

HERA-data in the light of small x evolution with state of the art NLO input

Janne Kuokkanen¹, Kari Rummukainen², and Heribert Weigert³

¹ Department of Physical Sciences, University of Oulu, P.O. Box 3000, FI-90014 Oulu, Finland

² Department of Physics and Helsinki Institute of Physics, P.O.Box 64, FI-00014 University of Helsinki, Finland

³ University of Cape Town, Dept. of Physics, Private Bag X3, Rondebosch 7701, South Africa

Both total and diffractive cross sections from HERA are successfully confronted with JIMWLK evolution equations in the asymptotic pseudo-scaling region. We present a consistent, simultaneous description of both types of cross sections that includes NLO corrections in the form of running coupling and energy conservation corrections. The inclusion of energy conservation corrections allows to match all available data with $x_{\text{bj}} \leq .02$ i.e. up to $Q^2 \leq 1200 \text{ GeV}^2$. We discuss the effects of quark masses including charm, contrast asymptotic and pre-asymptotic fit strategies, and survey non-perturbative uncertainties related to impact parameter dependence.

HIP-2011-23/TH

1 Introduction

Much of the abundant particle production in modern collider experiments at high energies is triggered by gluon channels, thus imprinting the features of gluon phase space on many observables. This is the basis of the importance of the Color Glass Condensate (CGC) [1–29] for virtually all current collider experiments be they designed to answer particle- or heavy-ion-physics questions.

The characteristic feature of enhanced gluon emission into the final state at high energies is the emergence of an energy dependent transverse correlation length $R_s(x)$. Its associated conjugate momentum scale $Q_s(x) \sim 1/R_s(x)$ signifies the onset of gluon saturation, hence the name saturation scale. As gluon numbers rise with energy, the correlation length of the dense gluon cloud shrinks, the saturation scale increases.

The mere existence of such an energy driven scale has led to a large body of phenomenological literature, often based on identifying quantities that can be expected to crucially depend on the saturation scale Q_s to gain insight into the energy dependence of some observable by applying a scaling argument. The origin of this idea predates the observation that perturbative QCD allows us to predict Q_s -scaling in the context of the JIMWLK equation (or in an independent scattering approximation the BK equation [21–25]) and has provided us with the discovery of what was called geometric scaling in HERA data [30–32]: plotting the ep cross sections measured at HERA not as a function of rapidity $Y = \ln(1/x)$ and momentum transfer Q^2 independently, but instead as a function of the scaling variable $Q^2/Q_s^2(x)$ reveals beautiful scaling features of the data for $x \leq 10^{-2}$ that extend even to diffractive measurements. This scaling subsumes the complete energy dependence of the data at $x \leq 10^{-2}$ in a single energy dependent function, $Q_s(x)$.

The standard approach to cross correlate HERA data in the Q^2 direction using Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution equations reproduces this scaling feature only accidentally, in the sense that they match data which themselves exhibit scaling. There is no *intrinsic* reason why such a scaling feature should emerge *within* the domain of validity of the linear DGLAP formalism¹. By contrast, in the CGC context, scaling is a natural byproduct of nonlinear features of gluon emission and saturation, which adds a particular interest to a comparison of CGC results with HERA data.

The derivation of the JIMWLK evolution equation relies on the prerequisite that there exists a frame in which the gluon field of the target becomes strong, while the gluon field of the projectile is weak. The resummation of the large target field then induces nonlinearities that capture the effects of gluon saturation and, through evolution, driven by perturbative gluon emission in the weak field projectile, induce saturation and the appearance of a saturation scale $Q_s(x)$ (its inverse $R_s \sim 1/Q_s$ has the interpretation of a transverse correlation length in the dense gluon cloud).

While the mere presence of such an energy dependent scale will impose its mark on any observable, strict mathematical scaling, i.e. the notion that *all* the cross sections and correlators of the theory share an energy dependence that can solely be expressed in term of that of the saturation scale $Q_s(x)$ is too naive an expectation: It is more natural that different n-point functions have their own associated scales unless tied together by JIMWLK evolution by belonging to the same Balitsky-hierarchy for which a strong factorization feature holds that allows to express these generically independent correlators to be expressed in terms of a single one, at least to good approximation. The scaling observed in the HERA total cross section should be interpreted as a signal for such a correlator factorization and the possibility to indeed truncate the Balitsky hierarchy of the dipole amplitude, the correlator dominating the total cross section.

Since the same correlators dominate both total and diffractive cross sections at HERA, the same x -dependent scale should also be in effect there, a fact established early in the Golec-Biernat–Wüsthoff (GB-W) model [30–32].

A second caveat arises from the QCD scale anomaly: even for observables which are predominantly determined by only a single correlator, true scaling is only to be expected at leading order (LO). At next to leading order (NLO) exact scaling receives *small* corrections [which manifest themselves as a slow drift of correlator shapes] purely induced by the scale anomaly, i.e. the running of the coupling. We will refer to this regime as the asymptotic or (pseudo-) scaling regime.

Despite the phenomenological success of models that assume (pseudo-) scaling of the cross section, it is by no means clear that the data would show clear evidence of strict or near scaling in HERA data or even for the HERA total cross section alone.

The question if the scaling observed in the γ^*p cross section [34–36] and the rapidity gap events [37–40] (see [41] for the most recent combined update of all ZEUS data) at HERA affects all correlators, i.e. represents true (pseudo-) scaling, or only signifies the presence of an intrinsic scale, indicative merely of the presence of nonlinearities, with scaling on a correlator level only apparent and potentially limited to a very narrow kinematic range is surprisingly hard to settle.

Let us emphasize that a survey of the literature presents us with a very ambiguous picture: All models from GB-W to the BK inspired parametrization of the dipole cross section by Iancu, Itakura

¹The derivation of scaling “within” DGLAP in [33] skirts the region of applicability of the argument: the scaling solutions shown there push into the region of large gluon densities where nonlinear/higher twist corrections are bound to become important.

and Munier [42] (IIM) to the BFKL+saturation boundary model of Mueller and Triantafyllopoulos [43, 44] (MT) exhibit either strict or pseudo-scaling of correlators, since scaling in this sense is one of the main constraining features in the construction of such models.

However, the shapes of these various (pseudo-) scaling solutions are notably different from each other and, as we will see below (see Fig. 7), from the scaling solution imposed by the evolution equation at NLO: the imposed scaling shapes in these models do *not* resemble the scaling shapes that emerge as solutions of the evolution equation.

Moreover, even within a CGC framework, strict or near scaling of correlators does not seem to be required to obtain a successful fit to the HERA total cross section: This was shown by Albacete, Armesto, Milhano, and Salgado in [45, 46]. They omit all NLO contributions beyond running coupling, but start evolution with a shape close to the GB-W parametrization. In this treatment, correlators only approach the scaling shape imposed by the evolution equation at the far end of the x range covered by the HERA experiments.

Our own simulations [47–49] add in energy conservation corrections to cover NLO corrections beyond the running coupling contributions and provide an equally convincing fit in the pseudo-scaling region. These same NLO corrections restrict our treatment to the region near (pseudo-) scaling due to stability and self consistency considerations that only emerge once all these NLO corrections are taken into account (see Sec. 2.3.3).

We will attempt to summarize the status of the theoretical tools presently available (state of the art are NLO evolution combined with LO impact factors, for a more in depth discussion see below) and explore what kind of tension available data pose on our theoretical analysis by considering both the HERA total cross section and the rapidity gap events.

As our first set of results emerges from fits to the total cross section: we will argue that simulations that include NLO corrections in the form of running coupling corrections and energy conservation corrections (the most complete set of NLO corrections currently available) favor fits in the pseudo-scaling region based on fits of the total cross section. The treatment of quark masses has proven to be somewhat problematic in earlier fits, which have generically used constituent like values of around 140 MeV for light quarks and 1.4 GeV for charm. Conceptually, quark masses are subleading in the small x limit where factors of the form $(\alpha_s)^{n+m}(\ln(1/x))^n$ are used to sort contributions by importance and hence should not pose a serious difficulty. Choosing current quark masses instead of constituent quark masses we find that fits are indeed feasible with the quark masses having their largest effect in the nonperturbative range with $Q^2 < 1 \text{ GeV}^2$. We find that the idea of [30] to address the situation by replacing x by some x_{eff} to modify the small Q^2 limit can improve the fit, although the specific form introduced in [30] proves unusable. Such resummations are by construction nonperturbative and should at this stage be taken to merely indicate the relevance of nonperturbative input in this range of phase space.

A second set of insights emerge from a study of rapidity gap events: Our analysis is based on the fit parameters extracted from total cross sections and clearly shows us the limits of a fit with incomplete NLO input, despite the quite satisfactory fit quality. A comparison with data clearly requires a $q\bar{q}g$ -component in the impact factor for which at present we only have a rough substitute which is only reliable in the large Q^2 limit and was already devised long ago in the context of the GB-W model [31]. This prevents us from using observables beyond the total cross section to get a closer look at the details of JIMWLK evolution – this crude treatment is not suitable for such a precision study. It turns out that non-perturbative aspects which enter through the impact

parameter dependence adds additional uncertainties as one steps beyond the total cross section. It affects the relative normalization of individual Fock-space components as soon as NLO impact factors start to play an important role in resolving the structure of the cross section.

This leads us to conclude that to address questions such as precision fits on initial conditions, the reliability of truncations of the full JIMWLK framework, or the size and nature of subleading N_c -corrections as advocated in [50] with any definiteness, we need full knowledge of all NLO contributions including the appropriate impact factors for the observable in question as well as an improved understanding of the non-perturbative aspects of the impact parameter dependence.

Any progress in this respect will also improve the utility of HERA fits as an input for fits to RHIC and LHC data and will be one of the most important tasks for the near future.

The structure of the paper is as follows: We begin by recapitulating the theoretical ingredients necessary to define the underlying observables at zeroth order in Sec. 2.1 and 2.2, putting some emphasis on the approximations underlying the expressions usually given in the literature. We review our present knowledge of the NLO corrections and how to use truncations of JIMWLK evolution to efficiently implement them in Sec. 2.3.

Sec. 3 is devoted to the systematics of a fit to the total cross section in the asymptotic pseudo-scaling region. We discuss general features of the asymptotic fit such as evolution speeds, Q_s in HERA phase space and correlator properties in Sec. 3.1. This is followed by a thorough study of the role of the energy conservation correction (Sec. 3.2), and the effect of quark masses (Sec. 3.3).

The fit obtained in Sec. 3 is then applied to diffractive data in Sec. 4. We begin our discussion in Sec. 4.2 with estimates of the non-perturbative uncertainties induced by our lack of knowledge of impact parameter- (b -) dependence of the eikonal correlators in a proton or nuclear target.

In Sec. 5 we compare our asymptotic approach that includes both running coupling corrections and the energy conservation correction with more conventional fits that only use the running coupling effect, but leave out the energy conservation correction and use the pre-asymptotic regime of small x evolution. We show that features of the solutions that are recovered perturbatively in the asymptotic region must be imprinted at least partially via the initial condition in the pre-asymptotic approach.

Sec. 5 collects our main results in attempt to provide a synthesis.

Several appendices provide a number of generic expressions (Sec. A), ancillary results partly indispensable to reconstruct our numerical simulations (Sec. B through D), as well as number of consistency checks (Sec. E).

2 Deep inelastic scattering at small x

2.1 Total γ^*A cross sections at zeroth order

There are many phenomenological applications to various physical processes and differential cross sections that are based on the idea of obtaining energy dependence from the scaling behavior in terms of the saturation scale. However, the observable directly addressed by JIMWLK evolution is simply the total cross section in very asymmetric collisions such as eA or pA experiments.

The strong asymmetry between projectile and target serves to justify the notion that the gluon field of the nuclear target with atomic number A can be thought of as much larger than that of

the much simpler electron or proton projectile. The asymmetry is used to describe the projectile in terms of a very simple wave function with only a very few (valence) partons, which then scatter on the large target field with a very large longitudinal momentum. This justifies the description of the interaction of such a projectile constituent with the large target field in a no recoil approximation. The no recoil approximation fixes the projectile constituent onto a worldline at a transverse position unaltered during the interaction with the target. Multiple interactions with the target field then build a non-Abelian eikonal factor, a path ordered exponential that captures the interaction of the constituent with the target field. This way the interaction of a quark in the projectile with the target is represented as an $SU(N_c)$ -valued field in the transverse plane U_x , an antiquark analogously enters as a U_x^\dagger and a gluon in the projectile interacts as \tilde{U}_x (the tilde denotes the adjoint representation). The graphical notation used to represent this is shown in Fig. 1 for the example of the zeroth order contribution to the $\gamma^* A$ scattering amplitude.

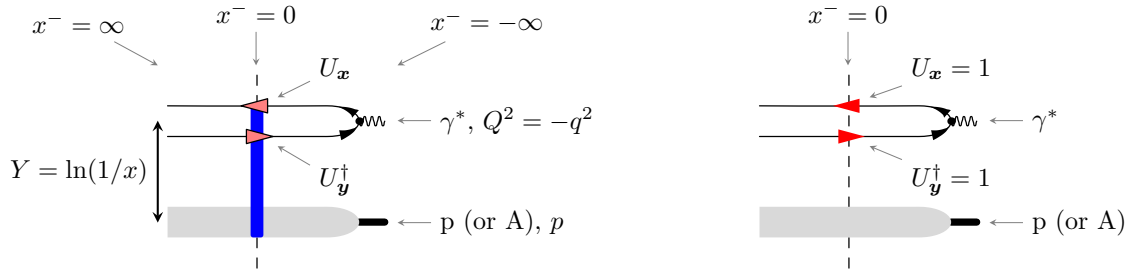


Fig. 1: Diagrammatic representation of the amplitude for $\gamma^* A$ scattering at small x at momentum transfer $Q^2 = -q^2$. Light cone “time” x^- runs from right to left. The interacting “out-state” (left) contain nontrivial interaction between projectile and target, which is marked by a vertical bar (blue online) at $x^- = 0$ that indicates the interaction region and markers for the Wilson lines picked up by the projectile constituents. The non-interacting “in-state” (right) instead has no interactions and correspondingly trivial Wilson line factors at $x^- = 0$.

The corresponding zeroth order total cross section arises as the absolute value squared of the difference of this interacting diagram with its noninteracting counterpart²

$$\left| \text{Diagram} - \text{Diagram} \right|^2 = \left(\text{Diagram} - \text{Diagram} \right) \left(\text{Diagram} - \text{Diagram} \right)^* . \quad (2.1a)$$

If the transverse momentum integrals are unrestricted, they will identify the transverse coordinates left and right of the cut, so that the the U -content of these diagrams is partially simplified:

$$(2.1a) = \text{tr} U_{y'} U_{x'}^\dagger U_x U_y^\dagger \xrightarrow[y' \rightarrow y]{x' \rightarrow x} \text{tr} 1 - \text{tr} U_x U_y^\dagger - \text{tr} U_y U_x^\dagger + \text{tr} 1 \quad (2.1b)$$

²Note that we keep the dashed vertical line that marks $x^- = 0$ also in the noninteracting case. We will need this below to distinguish where gluon vertices connect with respect to $x^- = 0$.

Note in particular that, with this assumption, the U -factors left and right of the cut in the $\overline{\text{out-out}}$ overlap cancel against each other. Diagrammatically, we have

$$(2.2)$$

While this is assumed universally in the literature without discussion, one should be clear that this is an approximation: phase space integrals over the transverse momenta \mathbf{k}_i of the final state quarks are limited by $W^2 := s_{\gamma^*A}$ (the invariant mass of the γ^*A -system) both in $t := -(\mathbf{k}_1 + \mathbf{k}_2)^2$ and the invariant mass of the produced pair. Consequently, also the momentum integrals over the transverse momenta are *not* unrestricted and thus the primed and unprimed coordinates are not *quite* the same in an exact treatment: even the total cross section will have a contribution from a four Wilson line operator. At small Q_s^2/W^2 , and presently achievable theoretical accuracy (see our discussion of NLO contributions below and the discussion of coincidence limits given in [50, 51]), it is fully justified to follow custom and take the limit shown in (2.2) and to incorporate this four point function only in a coincidence limit in which it becomes trivial.

With this caveat, it is the U -content of the two remaining diagrams, namely

$$\hat{S}_{\mathbf{x}\mathbf{y}}^{q\bar{q}} := \frac{\text{tr}(U_{\mathbf{x}}U_{\mathbf{y}}^\dagger)}{N_c} \quad (2.3)$$

and its complex conjugate that capture the interaction of the $q\bar{q}$ pair with the target.

The presence of the target wave function induces an energy dependent averaging process, that for the applications below will be strictly real, so that we may (and will) not distinguish $S_Y^{q\bar{q}}(\mathbf{x}, \mathbf{y}) := \langle \hat{S}_{\mathbf{x}\mathbf{y}}^{q\bar{q}} \rangle(Y)$ from its complex conjugate $\langle \hat{S}_{\mathbf{y}\mathbf{x}}^{q\bar{q}} \rangle(Y)$ in the following. For the total cross section, the target interaction of Eq. (2.1) can be fully summarized by the dipole operator

$$\hat{N}_{\mathbf{x}\mathbf{y}}^{q\bar{q}} := \frac{1}{N_c} \text{tr}(1 - U_{\mathbf{x}}U_{\mathbf{y}}^\dagger) \quad (2.4)$$

(and its complex conjugate), and the average $N_Y^{q\bar{q}}(\mathbf{x}, \mathbf{y}) := \langle \hat{N}_{\mathbf{x}\mathbf{y}}^{q\bar{q}} \rangle(Y)$, the dipole amplitude. It is the average over the target wave function, an operation that involves both perturbative and non-perturbative information, that proves the most difficult part of this calculation and induces the energy (or Y -) dependence of the cross section. The tool to extract this energy dependence is the JIMWLK equation, and its associated framework. At zeroth and leading order (LO) in $\alpha_s \ln(1/x)$ this allows us to describe the γ^*A cross section at a given energy entirely in terms of this simple amplitude without reference of “higher” Fock-space components of the projectile, simply by subsuming (in the sense of a renormalization group procedure) all other strongly interacting components into the averaging procedure.

The last ingredient of Eq. (2.1) not yet spelled out analytically, the wave function of the virtual photon, is known exactly [see for example [52] and Eqs. (2.6)], with both longitudinal and transverse polarizations contributing additively to the total cross section, $\sigma_{\text{tot}}^{\gamma^*p}(Y, Q^2) = \sigma_T^{\gamma^*p}(Y, Q^2) + \sigma_L^{\gamma^*p}(Y, Q^2)$.

This leads to an expression for the cross section in the form of a convolution in terms of the transverse coordinates which characterize the eikonal scattering position of the $q\bar{q}$ pair. Using

$\mathbf{r} = \mathbf{x} - \mathbf{y}$ and $\mathbf{b} = z\mathbf{x} + \bar{z}\mathbf{y}$ to denote dipole size and impact parameter respectively, one obtains (notations inspired by [53])

$$\sigma_{T,L}^{\gamma^*p}(Y, Q^2) = \sum_f \int d^2\mathbf{r} \int_0^1 dz \Phi_{T,L}^f(z, \mathbf{r}, \mathbf{r}', Q^2) \int d^2\mathbf{b} 2 N_Y^{q\bar{q}}(\mathbf{x}, \mathbf{y}) , \quad (2.5)$$

where z and $\bar{z} := 1 - z$ denote the longitudinal momentum fractions carried by the quark and antiquark respectively. For a fixed Q^2 , polarization and flavor f , the photon wave function product $\Phi_{T,L}^f(z, \mathbf{r}, \mathbf{r}', Q^2)$ ³ encodes the probability to find a $q\bar{q}$ pair of size $|\mathbf{r}|$, polarization T or L , and longitudinal momentum fraction z inside the virtual photon:

$$\begin{aligned} \Phi_T^f(z, \mathbf{r}, \mathbf{r}'; Q^2) = & \frac{\alpha_{em} N_c}{2\pi^2} e_f^2 \left((z^2 + \bar{z}^2) Q_f^2 \frac{\mathbf{r} \cdot \mathbf{r}'}{|\mathbf{r}| |\mathbf{r}'|} K_1(Q_f |\mathbf{r}|) K_1(Q_f |\mathbf{r}'|) \right. \\ & \left. + m_f^2 K_0(Q_f |\mathbf{r}|) K_0(Q_f |\mathbf{r}'|) \right) , \end{aligned} \quad (2.6a)$$

$$\Phi_L^f(z, \mathbf{r}, \mathbf{r}'; Q^2) = \frac{\alpha_{em} N_c}{2\pi^2} e_f^2 4Q^2 z^2 \bar{z}^2 K_0(Q_f |\mathbf{r}|) K_0(Q_f |\mathbf{r}'|) . \quad (2.6b)$$

In the above, e_f and m_f denote the charge and mass of the quark with flavor f and

$$Q_f^2 = z\bar{z}Q^2 + m_f^2 . \quad (2.7)$$

The impact parameter (b) integrated dipole amplitude has the interpretation of a $q\bar{q}$ -dipole cross section on the target:

$$\sigma_{q\bar{q}}(Y, (\mathbf{x} - \mathbf{y})^2) := 2 \int d^2\mathbf{b} N_Y^{q\bar{q}}(\mathbf{x}, \mathbf{y}) . \quad (2.8)$$

It carries the energy dependence of the cross section in terms of $1/x = e^Y$, the relative boost factor between the projectile and the target. The separation into wave-function factors (generically called impact factors) and dipole cross section (more generically Wilson line n -point functions) is prototypical to all observables in the high energy limit and extends to higher order in perturbation theory.

A description in terms of structure functions, F_2 , $F_{T,L}$ corresponds to a purely kinematical reparametrization according to the standard relation

$$\sigma_{\text{tot}}^{\gamma^*p}(x_{\text{bj}}, Q^2) = \frac{4\pi^2}{Q^2} \alpha_{em} F_2(x_{\text{bj}}, Q^2) = \frac{4\pi^2}{Q^2} \alpha_{em} (F_T(x_{\text{bj}}, Q^2) + F_L(x_{\text{bj}}, Q^2)) . \quad (2.9)$$

Please note that this does not imply a general link with particle distributions outside the parton gas region with $Q^2 \gg Q_s^2(x)$ where a twist expansion becomes valid.

³The notation for the expressions shown here and in Eq. (2.6) anticipate the diffractive case. There the structure of the wave function overlaps remains the same, only the dipole sizes on both sides of the final state cut must be distinguished.

2.2 Rapidity gaps in γ^*A at zeroth order

Apart from the total cross section, large complementary data set is available for rapidity gap events, in which the virtual photon fragments into (predominantly) a $q\bar{q}$ -pair accompanied by a gluon shower (which then hadronizes before it reaches the detector) that remains well separated from the target fragmentation region by a large rapidity gap. The kinematical setting for rapidity gap events is sketched in Fig. 2.

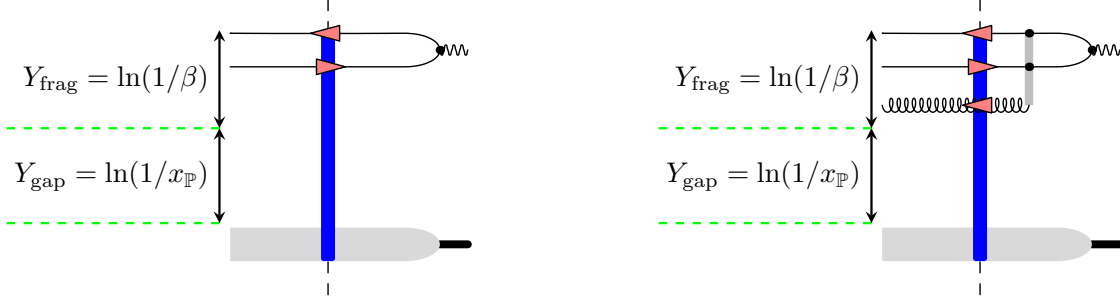


Fig. 2: Rapidity gap events differ from generic events contributing to the total cross section by a target side rapidity gap of size $Y_{\text{gap}} = \ln(1/x_{\mathbb{P}})$ into which no gluons are emitted. This gap is complemented by a projectile fragmentation range of size $Y_{\text{frag}} = \ln(1/\beta)$, such that $Y = Y_{\text{gap}} + Y_{\text{frag}}$.

The experimental situation is characterized by a quite strong similarity of the energy dependence of both total and diffractive cross sections.

Theoretically, all differences with the expression for the total cross section Eq. (2.1) arise from restrictions on the final state: With a rapidity gap on the target side, the target stays intact. Therefore, in the final state, the target is projected back onto its wave-function in each amplitude factor. This leads to separate target averages $\langle \dots \rangle(Y)$ in each amplitude. Since no net color can be exchanged across the gap, the $q\bar{q}$ -final state is necessarily projected onto a singlet. Once perturbative corrections are taken into account, and partons added to the projectile fragmentation region, they remain in an overall singlet.

At zeroth order in $\alpha_s \ln(1/x)$, the projectile only contains a $q\bar{q}$ pair, the cross section is given by

$$\left(\left(\text{projectile} - \text{target} \right) \supset \frac{1}{N_c} \subset \left(\text{projectile} - \text{target} \right) \right) \quad (2.10)$$

in close analogy with (2.1).

We note that each of these restrictions individually (the separate averages as well as the singlet projection of the projectile constituents) will prevent the simplification of the U -content that takes place in the $\overline{\text{out-out}}$ overlap for the total cross section due to Eq. (2.2). Instead

$$\left(\text{projectile} - \text{target} \right) \neq \left(\text{projectile} - \text{target} \right) \quad (2.11)$$

As an immediate consequence, not only the mixed ($\overline{\text{in}}$ -out) overlaps but also the $\overline{\text{out}}$ -out overlap will acquire nontrivial energy dependence.

What is left to understand are the restrictions on the transverse coordinates of Wilson lines in this expression as imposed by phase space integrals – again such a restriction alone would be sufficient to induce a nontrivial Wilson-line four-point function. To this end, note that β and Q^2 together determine the invariant mass of the projectile fragments:

$$\beta = \frac{Q^2}{Q^2 + M_X^2} . \quad (2.12)$$

This restricts the integration over transverse momenta. To be specific, assume n projectile constituents, denote longitudinal momentum fractions by z_i (with $\sum_i^n z_i = 1$) and final state transverse momenta by $\mathbf{k}_i := \tilde{\mathbf{k}}_i + z_i \mathbf{\Delta}$ (where $\mathbf{\Delta} := \sum_i^n \mathbf{k}_i$). Then the invariant mass of the n -particle final state is

$$M_{X,n}^2 = \sum_i^n \frac{\tilde{\mathbf{k}}_i^2 + m_i^2}{z_i} \quad (2.13)$$

so that a restriction on M_X imposes one linear constraint on the $\tilde{\mathbf{k}}_i^2$. To understand what this implies for the transverse coordinates, define

$$\mathbf{r}_i := \mathbf{x}_i - \mathbf{x}_n; \quad \mathbf{r}'_i := \mathbf{x}'_i - \mathbf{x}'_n; \quad \mathbf{b} := \sum_i z_i \mathbf{r}_i; \quad \mathbf{b}' := \sum_i z_i \mathbf{r}'_i; \quad (2.14)$$

and rewrite the exponents of the transverse momentum phase factors as

$$\sum_i^n \mathbf{k}_i \cdot (\mathbf{x}'_i - \mathbf{x}_i) = \mathbf{\Delta} \cdot (\mathbf{b}' - \mathbf{b}) + \sum_{i=1}^{n-1} \tilde{\mathbf{k}}_i \cdot (\mathbf{r}'_i - \mathbf{r}_i) . \quad (2.15)$$

This implies that integration over momentum transfer $t = -\mathbf{\Delta}^2$ will identify the light cone c.m. coordinates \mathbf{b} and \mathbf{b}' if one ignores the kinematical upper limit on t as for the total cross section. One of the $n - 1$ $\tilde{\mathbf{k}}_i$ integrations on the other hand is restricted by (2.13), so that one of the $n - 1$ independent distance pairings \mathbf{r}_i and \mathbf{r}'_i will remain independent after all unconstrained integrals are carried out. For the $q\bar{q}$ final state encountered at leading order, there is only one independent momentum variable available to begin with $\boldsymbol{\kappa} = \tilde{\mathbf{k}}_1 = -\tilde{\mathbf{k}}_2$. It is, therefore, tied directly to M_X via

$$M_{X,2}^2 = \frac{\boldsymbol{\kappa}^2 + m_i^2}{z\bar{z}} \quad (2.16)$$

and leaves behind phase factor containing $(\mathbf{r}' - \mathbf{r}) \cdot \boldsymbol{\kappa}$ where the dipole sizes \mathbf{r} and \mathbf{r}' in amplitude and complex conjugate amplitude remain independent. For each flavor independently, the length of $\boldsymbol{\kappa}$ is fixed in terms of β and the quark mass via (2.16) and (2.12) as

$$\boldsymbol{\kappa}_f^2 := z\bar{z}Q^2 \frac{\bar{\beta}}{\beta} - m_f^2 . \quad (2.17)$$

One obtains (see again [53])

$$\begin{aligned} \frac{d\sigma_{T,L}^{\gamma^* A \rightarrow Xp}}{d\beta}(x, Q^2) = & \frac{1}{\pi} \frac{Q^2}{4\beta^2} \sum_f \int d^2\mathbf{r} \int d^2\mathbf{r}' \int_0^{2\pi} \frac{d\varphi_\kappa}{2\pi} \int_0^1 dz \, z\bar{z} \, \Theta(\kappa_f^2) \, e^{i\kappa_f \cdot (\mathbf{r}' - \mathbf{r})} \\ & \times \Phi_{T,L}^f(z, \mathbf{r}, \mathbf{r}'; Q^2) \int d^2\mathbf{b} \, (N_Y^{q\bar{q}}(\mathbf{r}', \mathbf{b}'))^* N_Y^{q\bar{q}}(\mathbf{r}, \mathbf{b}) \end{aligned} \quad (2.18)$$

where $Y = \ln(1/x)$. Integrating this result over β leads to the total diffractive cross section. This step identifies \mathbf{r} with \mathbf{r}' if one extends the upper phase space boundary in M_X from W^2 to ∞ as discussed earlier. The result in turn maps back onto the total cross section if one removes the singlet projection in the final state which replaces $(N_Y^{q\bar{q}}(\mathbf{r}', \mathbf{b}'))^* N_Y^{q\bar{q}}(\mathbf{r}, \mathbf{b})$ by $(N_Y^{q\bar{q}}(\mathbf{r}', \mathbf{b}'))^* + N_Y^{q\bar{q}}(\mathbf{r}, \mathbf{b})$.

Note that this relationship of cross sections uniquely identifies the x arguments of the averages in the diffractive cross sections to be the overall Bjorken x of the process, not $x_{\mathbb{P}}$ as usually assumed in the literature. The numerical effect of such a replacement is, however, not large enough to affect the quality of any diffractive fits with state of the art expressions. These expressions suffer from more serious defects: incomplete NLO impact factors and nonperturbative normalization effects associated with the corresponding impact parameter averages as will be discussed below.

2.3 Beyond zeroth order

Leading order (LO) corrections to the above resum contributions proportional to $(\alpha_s \ln(1/x))^n$ and are fully taken into account by solving the LO-JIMWLK equation. The impact factors receive no corrections to their zeroth order form. At next to leading order, when contributions proportional to $\alpha_s(\alpha_s \ln(1/x))^n$ are taken into account, both JIMWLK evolution *and* the impact factors receive corrections.

Running coupling corrections to the evolution of Wilson line n -point functions have been calculated in full generality [54–56]. The remaining conformal corrections to evolution have been obtained by Balitsky and Chirilli [57], who presently work on the expression for NLO impact factors. A generalization of both aspects for arbitrary Wilson line n -point functions is yet to be devised. These ingredients would be required to extend the only existing treatment of JIMWLK-evolution [58] beyond leading order.

Fortunately truncations of JIMWLK evolution to finite sets of evolution equations, such as the Balitsky-Kovchegov (BK) equation or its more general GT (Gaussian truncation) counterpart allow us to use the available information to implement evolution at NLO. Accuracy on the impact factor side at present remains at LO. NLO accuracy for the impact factors introduces terms including $q\bar{q}g$ -correlators into the expressions for both the total cross section Eq. (2.9) and the diffractive cross section (2.18). We will see below that fits to the total cross section are not affected strongly, but that already the description of rapidity gap events suffers noticeably from this limitation.

Even with more modest goals in mind, such as the description of the total cross section it is mandatory to include NLO effects at least on the level of evolution equations. Here NLO corrections induce qualitatively new effects like scale breaking and a quantitatively important reduction in evolution speed compared to the LO situation. With the impact factors remaining at LO it is sufficient to know the Y -dependence of the dipole cross sections entering Eqs. (2.9) and (2.18),

but once the NLO impact factors are known this is no longer sufficient ($q\bar{q}g$ operators will require consistent treatment) and one is forced to either use the full JIMWLK-evolution framework or choose a “suitable” truncation. To appreciate what is involved, we briefly recapitulate the necessary tools.

2.3.1 JIMWLK and its truncations at LO

The JIMWLK evolution equation provides a means to calculate the energy- (or rapidity-) dependence of arbitrary U -correlators by first introducing an energy- (or rapidity-) dependent statistical weight $Z_Y[U]$ for the configurations of the U -fields. The dipole correlator of Eq. (2.4) is then expressed as a functional integral of the form

$$\frac{1}{N_c} \langle \text{tr}(1 - U_{\mathbf{x}} U_{\mathbf{y}}^\dagger) \rangle(Y) = \int \hat{D}[U] \frac{1}{N_c} \text{tr}(1 - U_{\mathbf{x}} U_{\mathbf{y}}^\dagger) Z_Y[U] \quad (2.19)$$

where $\hat{D}[U]$ is a functional Haar measure. This is meaningful in the sense that it allows to calculate the average of *any* operators, if it is possible to describe the evolution of *all* averages in terms of the evolution of the weight $Z_Y[U]$ defining the averaging procedure. This is the main content of the JIMWLK equation: it abstracts the energy dependence of the average from the operator being averaged by describing it as a functional evolution equation for $Z_Y[U]$. At LO, the equation takes the form of a functional Fokker-Planck equation

$$\frac{d}{dY} Z_Y[U] = -H_{\text{JIMWLK}} Z_Y[U] \quad (2.20)$$

that traces how additional gluons are added to the phase space of the projectile as one increases the energy of the collision. To arrive at (2.20) one needs to prove [18] that this equation indeed allows to find the energy dependence of arbitrary U -correlators, not just the simple $q\bar{q}$ -operator entering the dipole cross section. For such generic operators $\hat{O}[U]$, $Z_Y[U]$ defines a target average via

$$\langle \hat{O}[U] \rangle(Y) := \int \hat{D}[U] \hat{O}[U] Z_Y[U] \quad (2.21)$$

so that (2.20) implies an evolution equation for each such operator $\hat{O}[U]$ that takes the form

$$\frac{d}{dY} \langle \hat{O}[U] \rangle(Y) = -\langle H_{\text{JIMWLK}} \hat{O}[U] \rangle(Y) . \quad (2.22)$$

One of the features of JIMWLK evolution is that non-singlet correlators are exponentially suppressed by infrared divergent contributions, only singlet correlators survive.

A key feature of Y -evolution is that Eq. (2.20) gives rise to coupled hierarchies of evolution equations for U -correlators, known as Balitsky hierarchies. This already becomes manifest in the evolution equation for the $q\bar{q}$ operator at LO: It can be written as

$$\frac{d}{dY} \langle \text{tr}(U_{\mathbf{x}} U_{\mathbf{y}}^\dagger) \rangle(Y) = \frac{\alpha_s}{\pi^2} \int d^2 z \mathcal{K}_{\mathbf{x}\mathbf{z}\mathbf{y}} \left(\langle [\tilde{U}_{\mathbf{z}}]^{ab} \text{tr}(t^a U_{\mathbf{x}} t^b U_{\mathbf{y}}^\dagger) \rangle(Y) - C_f \langle \text{tr}(U_{\mathbf{x}} U_{\mathbf{y}}^\dagger) \rangle(Y) \right) \quad (2.23)$$

or, using (2.3) and the Fierz identity

$$[\tilde{U}_z]^{ab} 2\text{tr}(t^a U_x t^b U_y^\dagger) = \text{tr}(U_x U_z^\dagger) \text{tr}(U_z U_y^\dagger) - \frac{1}{N_c} \text{tr}(U_x U_y^\dagger) \quad (2.24)$$

as

$$\frac{d}{dY} \langle \hat{S}_{xy} \rangle(Y) = \frac{\alpha_s N_c}{2\pi^2} \int d^2 z \, \mathcal{K}_{xyz} \langle \hat{S}_{xz} \hat{S}_{zy} - \hat{S}_{xy} \rangle(Y) . \quad (2.25)$$

The integral kernel in both (2.23) and (2.25) is given by [3, 21]

$$\mathcal{K}_{xyz} := \frac{(\mathbf{x} - \mathbf{y})^2}{(\mathbf{x} - \mathbf{z})^2 (\mathbf{z} - \mathbf{y})^2} . \quad (2.26)$$

Eqs. (2.23) and (2.25) do not represent closed equations since the evolution of $\langle \text{tr}(U_x U_y^\dagger) \rangle(Y)$ depends on an operator with an additional gluon operator insertion made manifest by the \tilde{U} appearing in the first term on the right hand side of Eq. (2.23). The evolution equation of that new operator, $\langle [\tilde{U}_z]^{ab} \text{tr}(t^a U_x t^b U_y^\dagger) \rangle(Y)$, in turn will involve yet one more insertion of a gluon operator \tilde{U} , iteratively creating an infinite coupled hierarchy of evolution equations, the Balitsky hierarchy of the quark dipole operator (2.3) [23, 24]. JIMWLK evolution summarizes the totality of all such hierarchies, based on any (gauge invariant) combination of multipole operators but can only be solved numerically [51, 58, 59] at considerable computational cost.

One may note that within such hierarchies one finds numerous cross- and self-referencing patterns that can be exposed by looking at coincidence limits of coordinates in the operators involved. Taking $\langle [\tilde{U}_z]^{ab} \text{tr}(t^a U_x t^b U_y^\dagger) \rangle(Y)$ as an example, one finds its evolution equation linked with that of $q\bar{q}$ - and gg -dipole operator averages to which it reduces in the limits $\mathbf{x} \rightarrow \mathbf{y}$ and $\mathbf{z} \rightarrow \mathbf{x}$ (or \mathbf{y}) respectively. More generically, generalizing from $q\bar{q}$ dipoles to $\mathcal{R}\bar{\mathcal{R}}$ -dipoles (where \mathcal{R} refers to an arbitrary representation \mathcal{R}) one finds their evolution equation to contain the operator $[\tilde{U}_z]^{ab} \text{tr}(t^{\mathcal{R}^a} U_x t^{\mathcal{R}^b} U_y^\dagger)$ (see [51]). The evolution equations of this operator are mapped onto those of $\mathcal{R}\bar{\mathcal{R}}$ -dipoles or gg -dipoles in the two coincidence limits according to

$$\lim_{y \rightarrow x} [\tilde{U}_z]^{ab} \text{tr}(t^{\mathcal{R}^a} U_x t^{\mathcal{R}^b} U_y^\dagger) = C_{\mathcal{R}} \frac{d_{\mathcal{R}}}{d_A} \tilde{\text{tr}}(\tilde{U}_z \tilde{U}_x^\dagger) , \quad (2.27a)$$

$$\lim_{z \rightarrow y \text{ or } x} [\tilde{U}_z]^{ab} \text{tr}(t^{\mathcal{R}^a} U_x t^{\mathcal{R}^b} U_y^\dagger) = C_{\mathcal{R}} \text{tr}(U_x U_y^\dagger) . \quad (2.27b)$$

The more U -fields involved, the more constraints are imposed by coincidence limits and one may construct whole towers of operators linked downwards by coincidence limits. All these structures and relationships are automatically preserved and maintained in full JIMWLK evolution which, at least at one loop accuracy, can be simulated numerically.⁴

The theoretical picture can be simplified and the numerical effort required to solve the evolution equation can be reduced significantly by truncating the hierarchies. Any such truncation comes at

⁴Limitations are imposed only by available computational resources with limits on evolution ranges and initial conditions that can be accommodated without losing numerical accuracy.

the price of introducing an additional approximation. The most widely used truncation of JIMWLK evolution is known as the BK approximation. It assumes the factorization

$$\langle \hat{S}_{\mathbf{xz}} \hat{S}_{\mathbf{zy}} \rangle(Y) \rightarrow \langle \hat{S}_{\mathbf{xz}} \rangle(Y) \langle \hat{S}_{\mathbf{zy}} \rangle(Y) , \quad (2.28)$$

which turns Eq. (2.25) into a closed equation in terms of $\langle \hat{S}_{\mathbf{xzy}} \rangle(Y)$ only and thus decouples the rest of the Balitsky hierarchy. The BK truncation is valid and is parametrically justified in the large- N_c limit for scattering on a large dilute nuclear target. Using (2.28) in (2.25) we obtain the BK evolution equation

$$\frac{d}{dY} \langle \hat{S}_{\mathbf{xzy}} \rangle(Y) = \frac{\alpha_s N_c}{2\pi^2} \int d^2 z \, \mathcal{K}_{\mathbf{xzy}} \left[\langle \hat{S}_{\mathbf{xz}} \rangle(Y) \langle \hat{S}_{\mathbf{zy}} \rangle(Y) - \langle \hat{S}_{\mathbf{xzy}} \rangle(Y) \right] . \quad (2.29)$$

An alternative truncation has been discussed in [27, 51, 60] and dubbed the Gaussian truncation (GT) in [51]. In spirit, it approximates the JIMWLK average with Glauber-iterated two gluon t-channel exchange with the target.

This leads to explicit expressions for multi- U -correlators, in terms of a two point function $\mathcal{G}_{Y,\mathbf{xy}}$, for example

$$\langle \text{tr}(U_{\mathbf{x}}^{\mathcal{R}} U_{\mathbf{y}}^{\mathcal{R}\dagger}) \rangle(Y) = d_{\mathcal{R}} e^{-C_{\mathcal{R}} \mathcal{G}_{Y,\mathbf{xy}}} , \quad (2.30a)$$

$$\langle [\tilde{U}_{\mathbf{z}}]^{ab} \text{tr}(t^a U_{\mathbf{x}}^{\mathcal{R}} t^b U_{\mathbf{y}}^{\mathcal{R}\dagger}) \rangle(Y) = C_{\mathcal{R}} d_{\mathcal{R}} e^{-\frac{N_c}{2} (\mathcal{G}_{Y,\mathbf{xz}} + \mathcal{G}_{Y,\mathbf{zy}} - \mathcal{G}_{Y,\mathbf{xy}}) - C_{\mathcal{R}} \mathcal{G}_{Y,\mathbf{xy}}} \quad (2.30b)$$

and a consistent description of evolution for dipoles in all those arbitrary representations in terms of a single equation for \mathcal{G} (see [51])

$$\frac{d}{dY} \mathcal{G}_{Y,\mathbf{xy}} = \frac{\alpha_s}{\pi^2} \int d^2 z \, \mathcal{K}_{\mathbf{xzy}} \left(1 - e^{-\frac{N_c}{2} (\mathcal{G}_{Y,\mathbf{xz}} + \mathcal{G}_{Y,\mathbf{zy}} - \mathcal{G}_{Y,\mathbf{xy}})} \right) , \quad (2.31)$$

irrespective of the representation \mathcal{R} . One may think of GT as a truncation that, compared to BK, includes the minimal subset of $1/N_c$ suppressed contributions needed to restore the group theoretical coincidence limits (2.27) which are automatically satisfied by full JIMWLK evolution. It can be shown [51] that the dynamical content of GT and BK are in fact the same in the sense that replacing $q\bar{q}$ - and gg -dipoles appearing in Eqs. (2.30) (directly or in certain limits) by their large- N_c counterparts (proportional to $e^{-\frac{N_c}{2} \mathcal{G}_{Y,\mathbf{xy}}}$ and $e^{-N_c \mathcal{G}_{Y,\mathbf{xy}}}$ respectively) maps the BK equation (2.29) onto (2.31) and vice versa. GT improves over BK not in terms of dynamical content, but in the way this content is mapped onto different correlators of the theory.

It was shown in [51] at one loop accuracy, that the modification of the truncation encoded in Eqs. (2.30) and (2.31) under the name Gaussian truncation leads to slightly better agreement with full JIMWLK evolution than the BK truncation. We will use both BK and GT at NLO accuracy to compare to data below and will see that GT results in a slight improvement of the fit in keeping with the slightly better match of GT with JIMWLK evolution. The main advantage of the Gaussian truncation is that it allows to consistently describe general n -point functions such as that on the left of (2.2) before the local limit is taken. Unlike the BK approximation it is versatile enough to allow us to test the reliability of the phase space approximations that are built into (2.9) without being hampered by $\mathcal{O}(1/N_c^2)$ corrections.

2.3.2 Evolution at NLO

At NLO the picture gets even more complicated. At this accuracy not only single gluon Wilson lines are added to the original dipoles at the level of Eq. (2.25), also insertions of nonlocal operators such as $\text{tr}(t^a U_{\mathbf{z}_2} t^b U_{\mathbf{z}_2}^\dagger)$ (in the case of quark contributions) appear on the right hand side. It is clear that a full JIMWLK treatment becomes more and more costly and suitable truncations more and more of a necessity. In the light of the increasingly complicated insertions one would expect a strict leading $1/N_c$ BK approximation to become rather crude. The Gaussian truncation on the other hand should remain a viable candidate for a useful truncation. Here we choose an NLO treatment that, for the total cross section (2.9) and the diffractive cross section (2.18) allows us to use both BK and GT with comparable accuracy, and only keep in mind that once accuracy is high enough to consider testing the validity of (2.2) for the total cross section or more differential observables as discussed in [50], the Gaussian truncation becomes the tool to choose. Our main reason not to use the full dipole evolution as presented in [57] is numerical efficiency. While it is easy to include running coupling corrections according to [54–56], an implementation of the conformal corrections is impractical. Instead we adopt to substitute the conformal corrections with an energy (or rather longitudinal momentum) conservation correction as suggested by Gotsman, Levin, Maor and Naftali [61]. This is an attempt to resum DGLAP type corrections that enter small- x evolution at NLO that resum collinear contributions to all orders but should not lead to any double counting conflicts with the resummation of running coupling corrections.

In the BK-truncation the NLO equation to solve takes the form

$$\frac{d}{dY} S_{Y;\mathbf{x}\mathbf{y}} = \frac{N_c}{2\pi^2} \int d^2 z \mathcal{M}_{\mathbf{x}\mathbf{z}\mathbf{y}} \left(1 - \frac{d}{dY}\right) (S_{Y;\mathbf{x}\mathbf{z}} S_{Y;\mathbf{z}\mathbf{y}} - S_{Y;\mathbf{x}\mathbf{y}}) \quad (2.32)$$

while its GT counterpart reads

$$\frac{d}{dY} e^{-C_f \mathcal{G}_{Y;\mathbf{x}\mathbf{y}}} = \frac{N_c}{2\pi^2} \int d^2 z \mathcal{M}_{\mathbf{x}\mathbf{z}\mathbf{y}} \left(1 - \frac{d}{dY}\right) \left(1 - e^{-\frac{N_c}{2} (\mathcal{G}_{Y;\mathbf{x}\mathbf{z}} + \mathcal{G}_{Y;\mathbf{z}\mathbf{y}} - \mathcal{G}_{Y;\mathbf{x}\mathbf{y}})}\right) e^{-C_f \mathcal{G}_{Y;\mathbf{x}\mathbf{y}}} . \quad (2.33)$$

The kernel function $\mathcal{M}_{\mathbf{x}\mathbf{z}\mathbf{y}} \equiv \mathcal{K}_{\mathbf{x}\mathbf{z}\mathbf{y}} R_{\mathbf{x}\mathbf{z}\mathbf{y}}^{\text{eff}}$ is a product of the leading order BFKL/BK kernel $\mathcal{K}_{\mathbf{x}\mathbf{z}\mathbf{y}}$ of (2.26) and what one may call the effective running strong coupling $R_{\mathbf{x}\mathbf{z}\mathbf{y}}^{\text{eff}}$. The energy conservation corrections are represented by the derivative term on the right hand sides. Without it, both equations can be solved by a single step of numerical integrations based on the input of S_Y or \mathcal{G}_Y alone. To access the derivative one needs to know these functions at two Y values, Y and $Y + \Delta Y$, which forces us to use a much more costly iterative procedure described in App.C.

To arrive at a precise form for $R_{\mathbf{x}\mathbf{z}\mathbf{y}}^{\text{eff}}$ one should be aware that there exists no canonical way to separate running coupling corrections from the conformal contributions at NLO. Only their sum is unambiguously defined, to split them apart one is forced to introduce a separation scheme as discussed in [54–56, 62].

For data comparison we will adopt the separation scheme that subsumes most of the known NLO corrections into the running coupling contribution as suggested by Balitsky [56] instead of the scheme originally suggested in [54, 55]. We introduce the shorthand notations

$$r = |\mathbf{r}| = |\mathbf{x} - \mathbf{y}| , \quad r_1 = |\mathbf{r}_1| = |\mathbf{x} - \mathbf{z}| , \quad r_2 = |\mathbf{r}_2| = |\mathbf{y} - \mathbf{z}| , \quad (2.34)$$

to refer to coordinate differences⁵ and

$$\mu^2(r) = \frac{\mathcal{C}^2}{r^2}; \quad \mathcal{C}^2 = 4e^{-5/3-2\gamma_E}, \quad (2.35)$$

to refer to the scales in the coupling constants⁶. Now we may write the running coupling to be used in Eq. (2.32) as

$$R_{xyz}^{\text{eff}} = R^{\text{eff}}(r, r_1, r_2) = \alpha_s(\mu(r)) \left[1 + \frac{r_1^2}{r^2} \left(\frac{\alpha_s(\mu(r_2))}{\alpha_s(\mu(r_1))} - 1 \right) + \frac{r_2^2}{r^2} \left(\frac{\alpha_s(\mu(r_1))}{\alpha_s(\mu(r_2))} - 1 \right) \right], \quad (2.36)$$

which is a combination of three individual running couplings given by the standard perturbative one loop running

$$\alpha_s(\mu) = \frac{4\pi}{\beta_0} \frac{1}{\ln(\mu^2/\Lambda^2)}; \quad \beta_0 = (11N_c - 2N_f)/3; \quad N_f, N_c = 3 \Rightarrow \beta_0 = 9. \quad (2.37)$$

The choice to subsume as large a part of the NLO corrections into the running coupling has a large effect if all other NLO corrections are ignored. Comparing evolution the separation schemes of [54, 55] and [56] under these conditions has a large effect on the evolution speed as shown in Fig. 3, left panel. For completeness this figure also includes the effect of parent dipole running used widely in the literature. This procedure postulates an effective coupling $R_{\text{eff}} := \alpha_s(\mu(r))$ with the scale factor $\mathcal{C}^2 = 4$. This choice leads to almost the same evolution speed than Eq. (2.36) as can be seen from the leftmost of Figs. 3.

The idea behind the energy conservation correction is to include an all orders resummation of collinear corrections that start to contribute at NLO – they should make up the bulk of contributions not yet included. Unfortunately, we lack a derivation that would allow us to interlink its treatment with the separation schemes used to define the running coupling corrections: double counting of contributions can not be excluded without this information. Correspondingly, the energy conservation corrections do not alleviate the difference between evolution speeds induced by the different running coupling schemes once included in the calculation. This is shown in Fig. 3, middle panel and should be considered a major uncertainty in the present approach. However, this is of little consequence for data fits. Once we allow Λ to become a fit parameter, the difference of these treatments can be reabsorbed by a rescaling of that factor: after rescaling, the shapes of dipole amplitudes differ very little and allow fits of equal quality. We will find, however, that the energy conservation correction is a prerequisite to obtain a good fit in the pseudo-scaling region. A treatment that explicitly removes double counting from the outset would be much preferable.

One unavoidable complication remains in the region where $R_s(Y)\Lambda$ is near one, the region where the IR safety arguments that render JIMWLK evolution a self consistent procedure do not apply. JIMWLK evolution is justified where $R_s(Y)\Lambda \ll 1$, since at distances larger than the correlation length $R_s(Y)$, the correlator part of the evolution equations (2.32) or (2.33) very quickly approaches zero, so that for small enough R_s any sensitivity on how one regulates the Landau pole in the running coupling modified kernel \mathcal{M}_{xzy} disappears – IR uncertainties are effectively eliminated.

⁵ r is the size of the parent dipole ($q\bar{q}$ dipole) and r_1, r_2 refer to daughter dipoles (qg and $\bar{q}g$ dipoles). All expressions encountered are fully symmetric in interchange $r_1 \leftrightarrow r_2$.

⁶ $\gamma_E = 0.5772\dots$ is the Euler-Mascheroni constant. The scale factor $\mathcal{C}^2 = 4e^{-2\gamma_E-5/3}$ is specific to the $\overline{\text{MS}}$ scheme, it would be replaced by $4e^{-2\gamma_E}$ in the V-scheme.

If, however, $R_s(Y)\Lambda$ is near one, regulator effects become visible. They affect mostly the large r part of the dipole amplitude and integrated quantities such as the evolution speed, as defined in (2.39).

Since, as discussed below, the data lie in a range where the scale separation between Λ and Q_s is not safely established, we have to face uncertainties induced by our choice of regulator, our model for the IR behavior of the QCD coupling. The uncertainties encountered are shown in the panel on the right of Fig. 3 and involve an APT regulator [63, 64] and the extreme case of a practically unregulated coupling (corresponding to a cutoff treatment where α_s is the running coupling frozen only after it reaches 30.). This serves to illustrate the potential size and *range* of influence of the uncertainty induced by the presence of the Landau pole.

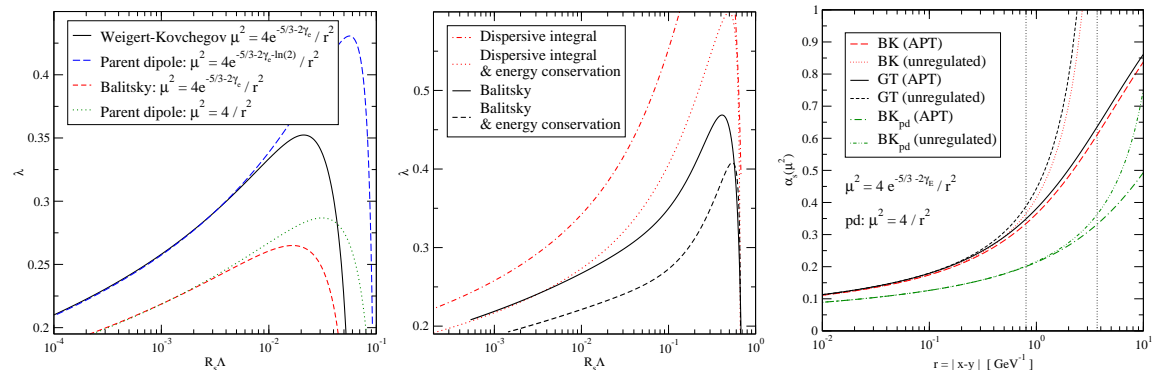


Fig. 3: **Left:** The evolution speed λ as a function of dimensionless units $R_s\Lambda$ for two different approaches mentioned in text. **Middle:** Evolution speed after adding in the energy conservation correction on top of running coupling **Right:** The running couplings (regulated vs. unregulated) used in different BK schemes plotted against the parent dipole size r . The vertical lines bracket approximately the region of saturation corresponding to HERA, $0.8 < R_s < 3.7$ for $5 \times 10^{-7} < x_{bj} < 0.02$.

2.3.3 Scaling behavior and self-consistency at NLO

JIMWLK evolution as well as both its BK and GT truncations are IR safe in the sense that in a transversely infinite medium and an initial condition with short enough initial correlation length $R_s(Y_0)$ all further evolution is governed by contributions on perturbative scales with most contributions arising near $R_s(Y) < R_s(Y_0)$. Under these conditions the integrand vanishes exponentially near the Landau pole and one may argue on physics grounds that contributions to evolution at the Landau pole may be neglected. In practice this is achieved by introducing some form of regulator. While at asymptotically high energies where $Q_s \gg 1$ GeV the precise choice of regulator *cannot* affect evolution, real data are far from that region and the choice of regulator will impact any comparison with experiment. For definiteness we choose what is known as an APT regulator, see for example [63, 64]⁷, and simply subtract the Landau pole from the expression for the running coupling. The advantage of using an APT regulator over the other possibilities is that it produces

⁷APT stands for Analytic Perturbation Theory. We use only its treatment of the Landau pole a practical way to regulate the running coupling.

a very smooth behavior of the effective running coupling R_{eff} even if it depends on all three scales μ_r , μ_{r_1} and μ_{r_2} . If one were to simply freeze the individual couplings below some fixed scale in R_{eff} of Eq. (2.36) the result would be discontinuous. This is partially compensated by a small \mathcal{K} in the evolution equation as long as energy conservation corrections are ignored. The stability of the iteration procedure used to implement the energy conservation correction described in App C, however, is strongly reduced by such discontinuities. The restriction to transversely infinite media is conceptually more serious: The JIMWLK framework does not correctly describe the transverse growth of a finite target. Near the edge, gluon densities become small and the evolution equations match up with the BFKL equation, unphysical Coulomb tails are no longer shielded by a density induced correlation length and nonperturbative contribution start to dominate. For this reason one usually uses JIMWLK equations to describe the in medium behavior and models any size effects. The simplest possible treatment for dipole cross section would be to postulate a separation of \mathbf{r} - and b -dependence in simple correlators like the dipole amplitude of Eq. (2.5) and introduce a profile function $T(\mathbf{b})$ to model the dipole cross section as

$$\sigma_{q\bar{q}}(Y, (\mathbf{x} - \mathbf{y})^2) = \sigma_0 N_{Y;\mathbf{x}\mathbf{y}}^{q\bar{q}} \quad \text{with} \quad N_{Y;\mathbf{x}\mathbf{y}}^{q\bar{q}} = N_{Y;\mathbf{x}-\mathbf{y}}^{q\bar{q}} \quad \text{and} \quad \sigma_0 := 2 \int d^2\mathbf{b} T(\mathbf{b}) . \quad (2.38)$$

With this assumption, all the nonperturbative information is encoded in a single constant σ_0 , which is then matched to data. While such a treatment is adequate for a LO treatment of the total cross section, we will see below that such an approach induces uncertainties already at leading order once more sophisticated observables are considered, rapidity gap events and the cross sections of [50] are among those. See Sec. 4.2 and the discussion leading up to it.

The numerical impact of NLO corrections is best illustrated by studying its impact on evolution speed λ , defined as

$$\lambda(Y) := -\frac{1}{\pi} \int \frac{d^2\mathbf{r}}{r^2} \frac{d}{dY} S_{Y;\mathbf{x}\mathbf{y}}^{q\bar{q}} . \quad (2.39)$$

This definition is equivalent to the “naive” definition of λ as the rate of change of the saturation scale, $\lambda(Y) = \frac{d}{dY} \ln(Q_s^2(Y))$, wherever strict scaling holds [65], but provides a useful generalization wherever strict scaling does not hold: This includes the pseudo-scaling case encountered as one steps beyond LO [58]. Where it becomes necessary to plot energy dependent quantities we often choose to plot against $R_s(Y)$ as the intrinsic Y -dependent scale of our simulations. For simplicity we take $R_s(Y)$ from

$$S_{Y;\mathbf{x}\mathbf{y}}^{q\bar{q}}(|\mathbf{r}| = R_s(Y)) = \frac{1}{2} . \quad (2.40)$$

Fig. 4 shows the behavior of the evolution speed as one incorporates NLO corrections: At leading order, i.e. with fixed coupling and without the energy conservation correction (Fig. 4, left) different initial conditions in the course of evolution towards higher Y (moving along the curves from right to left in the figures, from large $R_s(Y)$ to small $R_s(Y)$) all merge up with the asymptotic scaling regime in which λ becomes a constant. As running coupling is turned on (Fig. 4, middle), true scaling turns into pseudo-scaling: even the asymptotic speed remains Y -dependent. For $R_s(Y)\Lambda < 1$, evolution is slowed down drastically. Adding our last NLO ingredient (the energy conservation correction) leads to further slowdown as shown in Fig. 4 on the right. For comparison, all the panels in Fig. 4 show the λ range that leads to successful fits in our own fits and the scaling models of GB-W and IIM.

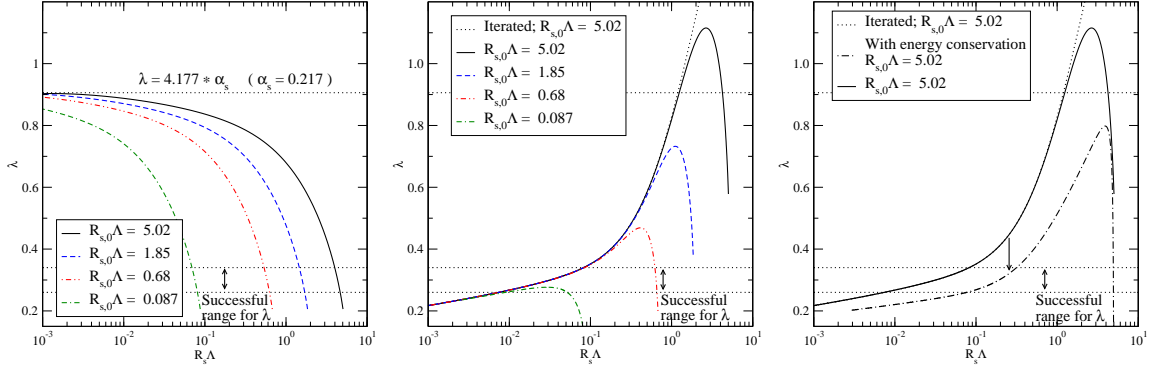


Fig. 4: **Left:** LO i.e. fixed coupling BK evolution: asymptotically the evolution speed $\lambda \rightarrow \text{const} \times \alpha_s$ for any relevant initial state. **Middle:** NLO BK evolution: the running coupling slows down the evolution. After the initial state effects are erased, the evolution speed settles on the same asymptotic line. **Right:** DGLAP type NLO corrections slow down the evolution further.

One concludes that NLO corrections to evolution (running coupling contributions in particular) are an essential ingredient to a successful fit to HERA data.

In preparation to the discussion of the fit procedure below let us note that we have intentionally plotted evolution speeds against $R_s(Y)\Lambda$, where, at this point, Λ is the QCD scale in the coupling. In the fits below, however, we will allow Λ to vary. The reason for this is twofold: we have observed already that such a change of scale can absorb most of the differences between different coupling separation schemes, beyond that one may also think of this a resummation of nonperturbative effects in the initial condition that affects both the values of $R_s(Y)$ at all Y and the associated evolution speeds $\lambda(Y)$ (through the size of the coupling). If the energy conservation correction is omitted as in [45], the latter is main justification to treat the scale factor as a fit parameter. In any case, it is through this fit procedure that we set the overall scale of $R_s(Y)$ in physical units.⁸

Evolution with NLO effects included allows us an additional cross check on the self consistency of our tools: For the calculation to be self consistent, there should be a clear hierarchy of size between leading and subleading contributions, unless the subleading contributions introduce a qualitatively new feature. This does apply to our discussion of evolution speed near the scaling regime: a new qualitative feature (scale breaking and the appearance of running coupling effects) change evolution speeds dramatically, while the effect of energy conservation corrections induce only minor corrections. This, however, is not universal: the relative size of the energy conservation corrections compared to evolution without it strongly depends on the shape of the solutions. Near or in the pseudo-scaling regime the energy conservation corrections have a relatively small effect. Far from the pseudo scaling regime, the energy conservation correction dominates the r.h.s of the evolution equation over a large range of $R_s(Y)$ scales. This difference in behavior is shown in Fig. 5. It is accompanied by severe numerical stability problems away from the pseudo-scaling region: We have employed iterative methods to obtain the derivative term as well as backwards difference methods (once two or more adjacent time steps are known with sufficient accuracy) and both require step sizes ΔY beyond anything even remotely practical to stabilize the numerical results⁹.

⁸ [45] fix the value of Λ and introduce a scale factor c to achieve the same goal.

⁹This is also the reason why we have chosen an approximate iteration procedure outlined in App. C to construct

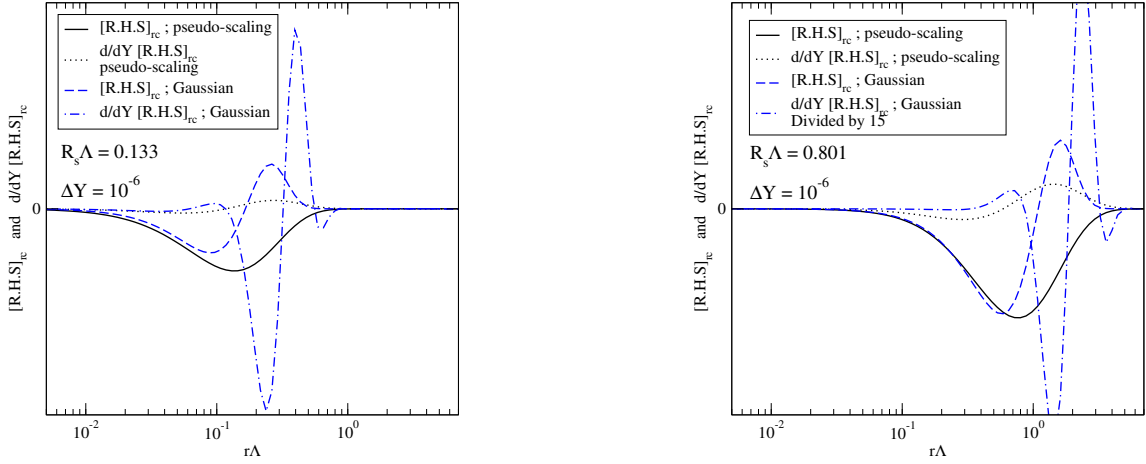


Fig. 5: Relative size of the energy conservation correction: the contributions to the r.h.s. of Eq. (2.33) at a fixed $R_s(Y)$ (left: $1/R_s \ll \Lambda$, right: R_s near Λ) are split up into the contribution without the energy conservation correction, labeled $[R.H.S.]_{rc}$ (it contains *only* the running coupling corrections) and the energy conservation correction $\frac{d}{dY}[R.H.S.]_{rc}$. In the scaling regime, the energy conservation correction is subleading. Away from the scaling regime (exemplified by a Gaussian correlator shape at the same correlation length) the energy conservation correction dominates.

Both these observations, the dominance of the energy conservation correction and the numerical stability, lead us to believe that the evolution equation in its present form is less reliable far away from the scaling regime. It would appear that additional resummations are necessary to reliably address the region far from scaling. The nature of such corrections is not known at present. We therefore advocate the use of the pseudo-scaling regime in data comparisons as long as fits in this region are at all possible.

3 Lessons from the total cross section

3.1 General features

In the following we will first confront the data with evolution assuming three light quarks with current quark masses $\lesssim 5$ MeV. This is sufficient to discuss the main features of the fit and any fit tensions. Fits with only a single mass are much faster to do and hence a more efficient tool to clarify such systematic questions.

After scouting the terrain in this manner we verify that the inclusion of quark masses will not change our conclusions and explore which phase space ranges are most affected by the inclusion of mass effects.

One generic feature is shared by *all* fits we have undertaken: the values of R_s appear to be well constrained by data, an impression of this is shown in Fig. 6 which shows alternatively $Q_s(Y) =$

the solution in the pseudo-scaling region.

$1/R_s(Y)$ and $Q_s(Y) = 4/R_s(Y)$ as they emerge from the fits overlaid on the data used. Sloped dotted lines indicate where the $q\bar{q}$ -dipole amplitude crosses 0.1 and 0.9 respectively. Details on these fits will be given below. With $1/R_s(Y)$ of the order of 1 GeV one would expect that nonperturbative contributions to evolution are inevitable, but to judge nonperturbative influence in a meaningful way, we need a more detailed analysis. Below we will comment on two aspects: IR effects in the evolution and phase space features in the case of mass effects.

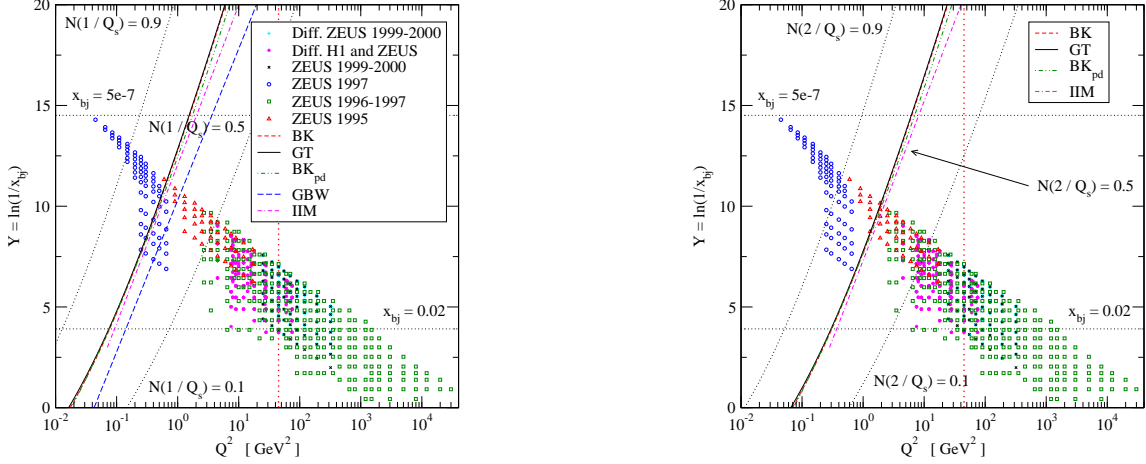


Fig. 6: **Left:** The phase space of all (inclusive and diffractive) HERA data included in fits together with the saturation momenta $Q_s^2(Y) = 1/R_s^2(Y)$ obtained from different fit approaches. The horizontal lines at $x = 5 \times 10^{-7}$ and $x = 0.02$ indicate the small x range used in the fits. In diffractive fits no restrictions were used. **Right:** To illustrate that slight changes of scale definitions strongly affect the appearance of this plot we replace $Q_s^2(Y) = 1/R_s^2(Y)$ by $Q_s^2(Y) = 4/R_s^2(Y)$.

The second feature that is quite well constrained by data, irrespective of details of the theoretical input, is the slope of $R_s(Y)$ in Fig. 6, or alternatively the evolution speed $\lambda(Y)$. Experience from GB-W and IIM model fits with $\lambda = .31$ and $\lambda = .29$ respectively as well as our fit experience with solutions from the evolution equations all establish a viable range for $\lambda(Y)$ that falls into the successful fit range indicated in Fig. 4. If this is generic, fits without energy conservation corrections will necessarily have smaller $R_s\Lambda$ values than fits that include such a contribution to evolution (Fig. 4, right). Since the physical values of $R_s(Y)$ are very well constrained this implies that fits without the energy conservation correction require a noticeably smaller value for Λ that fits that include energy conservation corrections. The former get the contributions from a region with smaller coupling and are less sensitive on the choice of IR regulator for the Landau pole. The considerably smaller Λ values necessary without the energy conservation corrections, however, would indicate a larger nonperturbative resummation entering the evolution equation: none of the alternatives is free of nonperturbative effects.

In an actual fit with solutions of an evolution equation, one needs to choose an initial condition, in our case always of the form of the GB-W model, specify the initial correlation length in units of Λ and evolve this initial condition in Y . As we proceed solving the equation the correlator shapes change away from the GB-W form and approach the asymptotic or pseudo-scaling regime. $R_s(Y)$ shrinks and λ traces a trajectory as indicated in Fig. 4. For any given trajectory one then needs to

relate the rapidity variable of the simulation to the physical rapidity by determining what we call Y_{offset} or Y_{off} to pinpoint $Y = 0$.¹⁰ On a given trajectory, Y_{off} is most closely linked to evolution speed. The fit of Y_{off} is done simultaneously with a fit of the physical units on $R_s(Y_0 = -\ln(2 \cdot 10^{-2}))$ by varying Λ , and a fit the overall normalization σ_0 from Eq. (2.38).

A fit in the pseudo-scaling region, where the shape of the dipole amplitude is fully determined by the nonlinearity with details of the initial condition erased everywhere but in the extreme UV, the fit is therefore a three parameter fit in terms of Y_{off} , Λ , and σ_0 .

Away from the asymptotic pseudo-scaling region many additional features of the initial condition survive and may affect the quality of the fit. At present we have no systematic tools to scan the space of initial conditions and very little constraints from theory. Our efforts below are meant to shed some first light on the issue using a very hands on attitude that is limited in scope mostly by the cost of creating a single trajectory on the one hand and the search of fit parameters for a given trajectory on the other.

3.2 Systematics from fits with light quarks, energy conservation included

Before we turn to fit quality, let us first collect the main features the data fits impose on correlator shapes and evolution speed in Fig. 7.

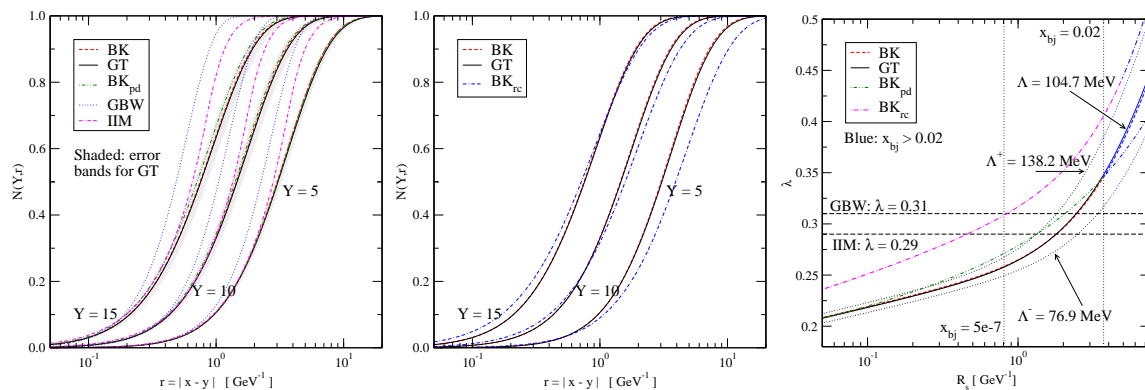


Fig. 7: **Left:** The evolution of $N_Y^{q\bar{q}}(r)$ in different approximations and models. The data align JIMWLK based descriptions and the IIM model in the region $N_Y^{q\bar{q}}(r) \lesssim 0.4$ – the GB-W model deviates significantly. **Middle:** The asymptotic (perturbatively imposed) solutions of the evolution equations are steeper after the energy conservation corrections are included. This benefits the fit quality overall *and* at large Q^2 in particular. **Right:** The evolution speed λ for all JIMWLK approximations considered as a function of correlation length R_s . Error bands are extracted from the condition $\chi^2/\text{dof} = 1$, see Appendix B.

We have already stressed at the outset of Sec. 3, the most tightly constrained feature in all the fits is $R_s(Y)$ in physical units. This also manifests itself in the fact that after the fit is performed, the shapes of dipole amplitudes agree very well within an order of magnitude below $R_s(Y)$ no matter

¹⁰Since we only start compare to data for $x \leq 2 \cdot 10^{-2}$ we may well obtain negative values of Y_{off} and still have dipole cross sections to cover all the data range considered. We have no ambitions to extend or parametrizations to larger x .

if the calculation is based on an evolution equation or a model. This is shown in Fig. 7, on the left panel for BK, GT, BK with parent dipole running (labeled BK_{pd}) and the two models GB-W and IIM. The models show noticeable deviations from the fits based on the pseudo-scaling solutions from evolution equations at $|\mathbf{r}| \gtrsim R_s(Y)$.

We had already seen in Fig. 4 that the main effect of the energy conservation correction is a further slowdown of evolution above that induced by the running of the coupling. The middle panel in Fig. 7 illustrates the second major impact by comparing BK and GT fits (which include the energy conservation correction) and an asymptotic BK fit without this correction labeled $[BK]_{rc}$: correlator shapes in the asymptotic regime stay *steeper* than in the corresponding evolution with the energy conservation correction omitted.

Evolution speeds corresponding to the fits are shown in the rightmost panel in Fig. 7. We indicate an associated error band only for the GT fit. It encloses all fits that include the energy conservation correction, but is clearly separated from the fit without the energy conservation corrections $[BK]_{rc}$, complementing the shape deviation already observed. The two models, by construction, have constant evolution speeds that fall near the average of evolution speeds obtained from the BK and GT fits, confirming our expectations.

Table 1 provides an assessment of the quality of the fits illustrated previously in terms of dipole amplitudes and evolution speeds in Fig. 7. The table reflects our fit strategy of first obtaining a fit at low $Q^2 \leq 45 \text{ GeV}^2$ which we then attempt to extend to a larger Q^2 range up to 1200 GeV^2 , monitoring any change in fit parameters required in the process. In the larger Q^2 range the models (GB-W and IIM) are known to fail, they were designed only with small Q^2 values in mind, but also the small x approximation employed in the evolution equations will have to break down eventually – even with partial resummations like the energy conservation correction built in. The success of such an extension is a measure as to how efficiently the resummations recapture large Q^2 effects.

$x_{bj} \leq 0.02$		BK_{pd}	BK	GT	GB-W	IIM	$[BK]_{rc}$
	$\lambda(x_{bj})$ $x_{bj} \in [\frac{5}{10^7}, 0.02]$	0.27-0.34	0.26-0.35	0.26-0.35	0.31	0.29	0.31-0.44
	$\Lambda \text{ [MeV]}$	$82.4^{+31.7}_{-24.4}$	$93.7^{+30.7}_{-25.1}$	$104.7^{+33.5}_{-27.8}$	$x_0 = \frac{1.1}{10^3}$	$x_0 = \frac{1.1}{10^4}$	50.4
$Q^2 \leq 45 \text{ GeV}^2$	$\chi^2/224$	0.818	0.811	0.810	1.401	0.828	1.760
	$\sigma_0 \text{ [GeV}^{-2}]$	54.01	55.05	55.33	44.59	51.47	56.84
$Q^2 \leq 1200 \text{ GeV}^2$	$\chi^2/295$	0.979	0.812	0.805	1.978	1.037	4.783
	$\sigma_0 \text{ [GeV}^{-2}]$	53.69	55.01	55.35	44.48	51.08	54.79

Table 1: Fit results to inclusive data from [34–36, 41]. χ^2/dof is below one for a wide range of Λ values indicated by the errors listed, see Fig. 19.

The best fits are obtained when we use the NLO evolution equations once energy conservation is included (columns labeled BK and GT as well as BK_{pd}):¹¹ The fit quality is excellent over both the small and the large Q^2 range. Already the low Q^2 range ($Q^2 < 45 \text{ GeV}^2$) determines both the physical units of $R_s(Y)$ and the ideal evolution speed $\lambda(Y)$ (via Y_{off}). To extend the fit to the full Q^2 range covered by the data below $x = 2 \cdot 10^{-2}$ only σ_0 needs to be readjusted. Note the *excellent*

¹¹All these fits are performed with light quarks only, the role of physical quark masses is discussed in Sec. 3.3.

χ^2/dof values over the whole Q^2 -range for both BK and GT truncations in the pseudo-scaling region.

An asymptotic fit with the energy conservation omitted is clearly unworkable: its χ^2 -value is barely acceptable already for $Q^2 \leq 45 \text{ GeV}^2$ and indicates an outright failure in the broader Q^2 region up to 1200 GeV^2 – it fares worse than the models (if extrapolated into this region) by far. The origin for this is the shallower shape of the dipole amplitudes observed in Fig. 7 (middle panel). The data require steeper correlators and lower evolution speeds. Evolution speeds do slow down as evolution proceeds to smaller $R_s(Y)\Lambda$, however, at the same time correlators flatten out further to asymptotically approach the shapes of the fixed coupling case as the running of the coupling slows down with shrinking $R_s(Y)\Lambda$ values. This leaves a very small window for Λ (viewed as a fit parameter) in which the correlators are still tolerably steep, but evolution speed is already small enough. As a result all features of this fit deviate from those shown in the left panel of Fig. 7: evolution speed is still quite large, and the match of the dipole correlators below R_s observed in the left panel of Fig. 7 is lost as well.

The success of the asymptotic fit with energy conservation included could be interpreted as an indication that the inclusion of the energy conservation correction gives a better match to the perturbative anomalous dimensions that govern the large Q^2 -behavior. This is in contrast to the preasymptotic fit of [45,46] (which omits energy conservation corrections) where relics of the steepness of the initial condition –non-perturbative in nature– allow for a good fit quality at all Q^2 . We will provide more details on this comparison in Sec. 5.

The models, GB-W and IIM, are clearly limited to a smaller Q^2 region. This holds even for the IIM model, which does incorporate additional perturbative information in the form of BFKL anomalous dimensions.

Fig. 8 shows asymptotic solutions and models against a subset of data to illustrate fit quality in different Q^2 ranges. The asymptotic fits with energy conservation included remain valid to astonishingly large Q^2 values, exceeding the 1200 GeV^2 range over which we have kept track of χ^2 -values above. This generic picture is reinforced if we contrast the scaling behavior of theory

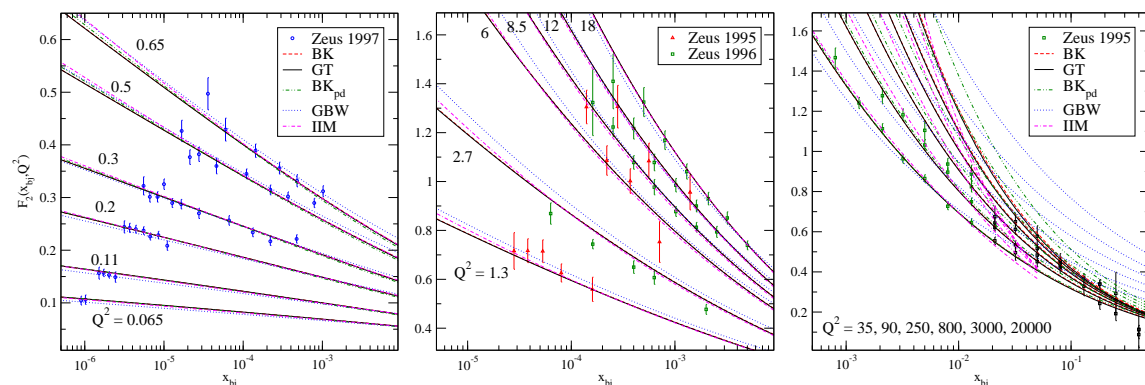


Fig. 8: $F_2(x, Q^2)$ as a function of x from a fit with light quarks compared to a subset of data. Different descriptions are qualitatively the same at low and moderate values of Q^2 (on the left and middle respectively). At high Q^2 (on the right) only the BK and GT fits compare well to the data.

and data as done in Fig. 9 for GT (left), BK (middle) GB-W fits (right). Shown are γ^*p cross

sections (see Eq. (2.9) for the connection $\sigma_{\gamma^*p} \leftrightarrow F_2$) as a function of scaling variable $Q^2/Q_s^2(x_{bj})$ for the three cases after the fit is performed. The borders between moderate and high virtualities (above $Q^2 = 120 \text{ GeV}^2$) are roughly indicated by the arrows overlaid on the plots. Data sets from different running periods are plotted by using different colors and symbols¹². To avoid overlapping, the fit results are separated from the experimental data by dividing the normalization factors σ_0 out. Note that the saturation scales are different functions of x_{bj} in each of the panels, although the differences between the two truncations (GT, left and BK, middle) are so small that no deviation can be discerned visually: both slopes coincide excellently with the experimental data up to the largest $Q^2/Q_s^2(x_{bj})$. Contrary to that, the GB-W model shown on the right can not resolve the high Q^2 data: the *slopes* of data and theoretical predictions start to deviate at large $Q^2/Q_s^2(x_{bj})$. A more detailed picture of the fit quality is given in Fig. 10.

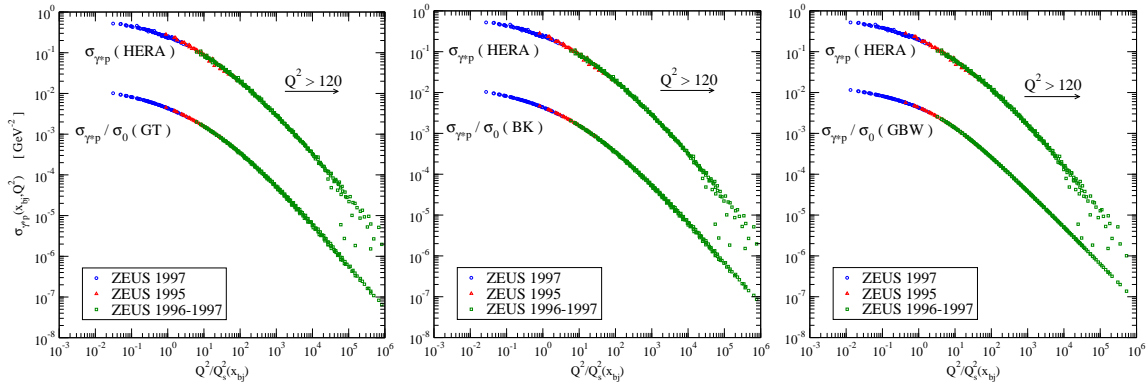


Fig. 9: The γ^*p cross section as a function of scaling variable Q^2/Q_s^2 . The plots show both data and theoretical predictions (shifted downwards by σ_0 to avoid clutter). **Left:** GT. **Middle:** BK. **Right:** The GB-W-model. Data from [34–36].

3.3 Quark masses

As already mentioned, quark masses are formally a subleading effect from the perspective of our small x resummations, but they do impact final state phase space and the width of wave functions in impact factors quite severely. The use of constituent quark masses of 140 MeV for light quarks and 1.4 GeV for charm is quite widespread in the context of the GB-W model and typically used in a restricted Q^2 -range (below 45 GeV^2). Since quark masses by nature are a nonperturbative feature that should have its main effect at small Q^2 one should expect that the inclusion of quark masses will not spoil the excellent fit quality that was obtained in our light quark fits for Q^2 above the heaviest quark included. From diffractive measurements we know that charm quarks should contribute significantly to the HERA cross sections while bottom quark contributions are negligible. One should therefore complement the three light quarks used above with a charm quark. This brings in a quark mass of 1.2–1.4 GeV depending on whether one considers current or constituent quarks, both of which are of the same order as $1/R_s(Y)$ in physical units and thus one would expect at least some complications in the nonperturbative sector. A straightforward fit with three light and one

¹²The same notation for the inclusive data is applied throughout.

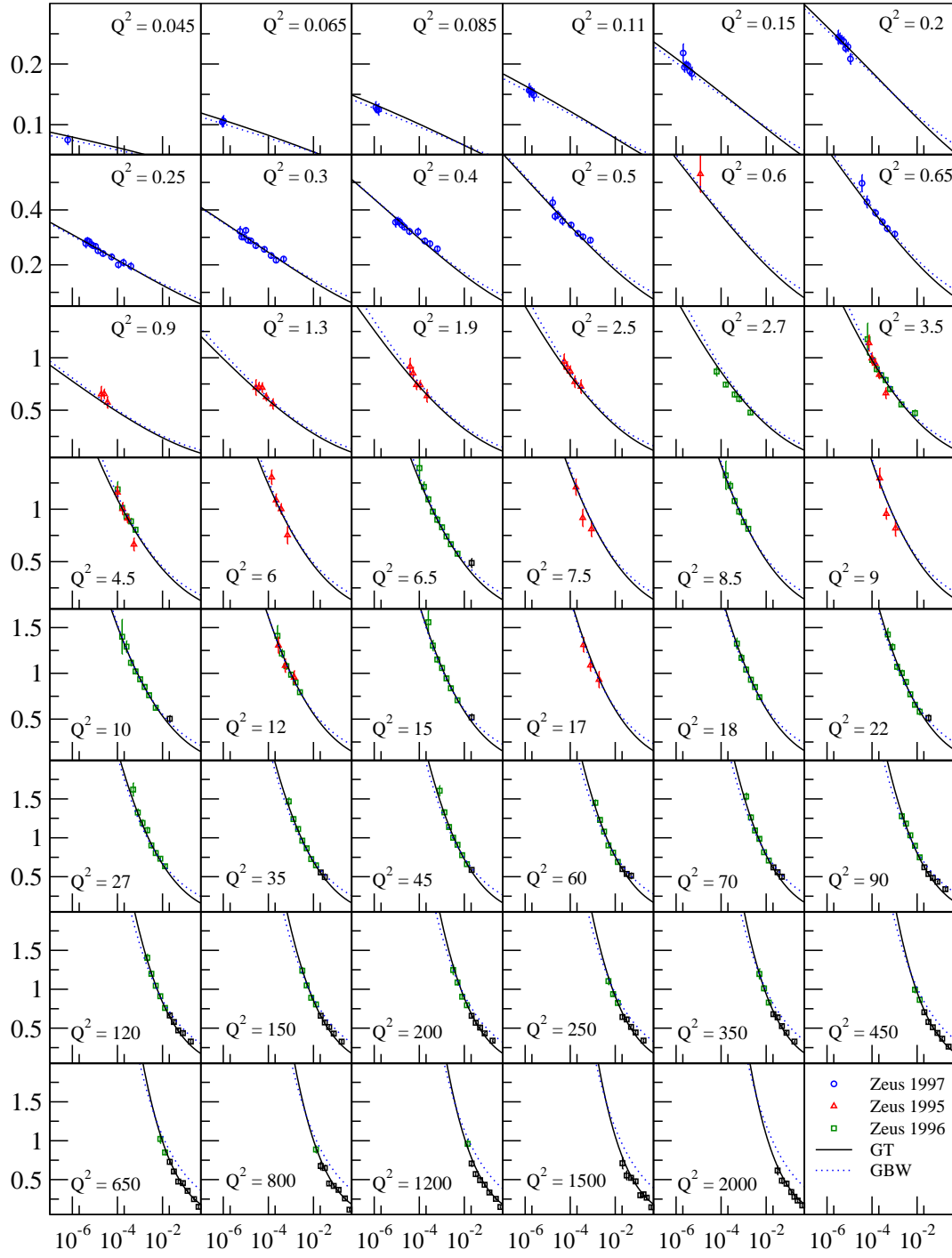


Fig. 10: $F_2(x_{bj}, Q^2)$ as a function of x_{bj} . The data are from ZEUS [34–36].

heavy quark shows that the inclusion of the charm quark reduces fit quality mainly in the Q^2 range below 1 GeV² (First two columns in Table 2). This is about as far as we can go in our analysis without introducing any model elements that modify the low Q^2 behavior in some ad hoc manner.

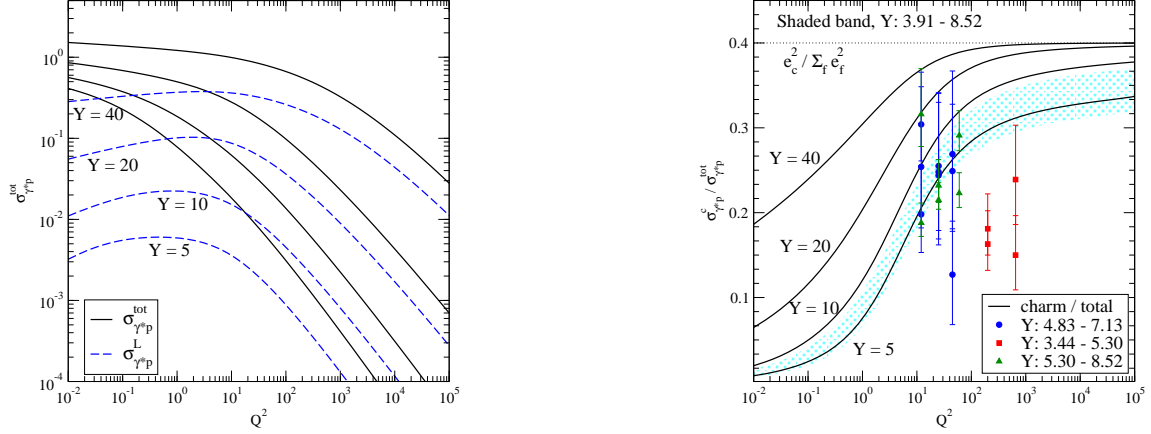


Fig. 11: Cross sections including charm quarks with $m_{u,d,s,c} = \{3, 5, 105, 1270\}$ MeV for $Y = \{5, 10, 20, 40\}$. The curves are calculated with the parameters corresponding to $x_{\text{eff}} = x(1 + Q_s^2/Q^2)$ (or equivalently $Y_{\text{eff}} = Y - \ln(1 + Q_s^2/Q^2)$) in Table 2. **Left:** $\sigma_{\gamma^*p}^{\text{tot}}$ and $\sigma_{\gamma^*p}^L$ (blue dashed) plotted separately. Below the largest quark mass used, σ^L starts to drop. This is incompatible with current conservation in the limit $Q^2 \rightarrow 0$ and reduces the quality of the fit. **Right:** The charm fraction $\sigma_{\gamma^*p}^c / \sigma_{\gamma^*p}^{\text{tot}}$ as a function of Q^2 . The horizontal line indicates the large Q^2 limit $e_c^2 / \sum_f e_f^2 = 2/5$. The data are from [66–68]. Agreement is clearly qualitative at best.

We should, however, at least qualitatively discuss a nonperturbative modification introduced by Golec-Biernat and Wüsthoff [30], a modeling device to accommodate nonperturbative contributions at small Q^2 . They have suggested to evaluate the dipole cross section at

$$x_{\text{eff}}^{\text{GB-W}} = x \frac{Q^2 + 4m_f^2}{Q^2} = \frac{Q^2 + 4m_f^2}{Q^2 + W^2} \quad (3.1)$$

instead of x in order to guarantee W -independent cross sections at small Q^2 as required by current conservation in the photo-production limit at $Q^2 = 0$. We have found that with the data set used here, (3.1) in fact improves the fit for $Q^2 \leq 45$ GeV² in the GB-W model. When using asymptotic solutions to the evolution equation, however, fit quality goes down (taken over the full Q^2 range) as shown in the third column of Table 2. In fact even the “inverse” modification $x_{\text{eff}} = x \frac{Q^2}{Q^2 + 4m_f^2}$ works better from a χ^2 perspective (fourth column), despite a clear lack of supportive arguments. One possible reason for the failure of (3.1) is conceptual: it is not compatible with the factorization into impact factors and Wilson line correlators that is at the core of our renormalization group picture. Quark masses are properties of projectile constituents and have every reason to show up in impact factors. On the other hand, it is hard to imagine that they should be resummed into a feature of the energy dependence of the Wilson line correlators, which are purely determined by target properties. But this is exactly what is done when using (3.1). From this perspective the only scale available for use in a modification at small Q^2 is in fact $Q_s^2(x)$. We tentatively suggest

to replace (3.1) with

$$x_{\text{eff}} = x \frac{Q^2 + Q_s^2(x)}{Q^2} = \frac{Q^2 + Q_s^2(x)}{Q^2 + W^2} . \quad (3.2)$$

While this does not freeze the cross sections at fixed W , it does flatten them noticeably at small Q^2 and leads to an improvement of the fit quality, both with a schematic mass pattern of $m_{u,d,s,c} = \{5, 5, 5, 1400\}$ MeV and the current quark mass pattern $m_{u,d,s,c} = \{3, 5, 105, 1270\}$ MeV as shown in the last two columns of Table 2. We emphasize that (3.2) is only a conjecture to resum part of the nonperturbative contributions below Q_s , one should not take too much encouragement from the improvement of χ^2 alone, see the ad hoc success of the redefinition in column 4.

x_{eff}	x	x	$x \frac{Q^2 + 4m_f^2}{Q^2}$	$x \frac{Q^2}{Q^2 + 4m_f^2}$	$x \frac{Q^2 + Q_s^2}{Q^2}$	$x \frac{Q^2 + Q_s^2}{Q^2}$
		$(Q^2 > 1 \text{ GeV}^2)$		ad hoc mod.		(phys. masses)
Y_{off}	−5.86	−6.87	−5.63	−5.99	−6.50	−6.73
Λ [MeV]	88.1	63.5	89.7	90.5	70.3	66.1
$\sigma_0 \text{ GeV}^{-2}$	57.48	73.04	61.20	51.69	68.94	70.52
χ^2/dof	1.41	1.05	1.59	1.23	1.13	1.14

Table 2: Including a charm quark; nonperturbative modifications. For the second column data with $Q^2 \leq 1 \text{ GeV}^2$ are removed, reducing the dof from 295 to 221. Quark masses are $m_{u,d,s,c} = \{5, 5, 5, 1400\}$ MeV in all but the rightmost column where $m_{u,d,s,c} = \{3, 5, 105, 1270\}$ MeV.

4 Lessons from the diffractive data

4.1 The need for NLO contributions to the impact factors

Diffractive HERA data extend down to $\beta \sim .04$. Thus overall $Y = \ln(1/x)$ and $Y_{\text{gap}} = \ln(1/x_{\mathbb{P}})$ remain comparable while $Y_{\text{frag}} = \ln(1/\beta)$ remains too small for multiple gluon emission to build up within the projectile fragmentation region – contributions from the $q\bar{q}g$ -component of the impact factor (which has its first contribution at NLO), however start to play a role even at such moderate β values as already observed in the pioneering papers of [69, 70] and reiterated in [71]. The main reason for that is that the $q\bar{q}$ contributions given in Eq. (2.18) strictly vanish at $\beta \rightarrow 0$, even with NLO effects to the evolution of the Wilson line correlators taken into account (see Figs. 13). Any gluon component in the impact factor, on the other hand, will generate a nonvanishing cross section in this region of phase space: the NLO contributions to the impact factors are the *leading* contribution at small β .

Unfortunately, no full expression for the $q\bar{q}g$ -component is available, only the large Q^2 and small β limits (without β -evolution) are known exactly. However, an interpolating form has been suggested in [53]. Below we find that nonperturbative contributions related to the target profiles by far dominate the uncertainties, and, for simplicity, we content ourselves with the large Q^2 expressions of [69, 70] to estimate the contributions.

The starting point then is

$$\frac{d\sigma^{\gamma^* A \rightarrow Xp}}{d\beta} = \frac{d\sigma_{q\bar{q},T}^{\gamma^* A \rightarrow Xp}}{d\beta} + \frac{d\sigma_{q\bar{q},L}^{\gamma^* A \rightarrow Xp}}{d\beta} + \frac{d\sigma_{q\bar{q}g,T}^{\gamma^* A \rightarrow Xp}}{d\beta} \Big|_{\text{LL}(Q^2)} \quad (4.1)$$

where the corresponding structure functions $x_{\mathbb{P}} F_{q\bar{q},T}^{D(3)}$, $x_{\mathbb{P}} F_{q\bar{q},L}^{D(3)}$, and $x_{\mathbb{P}} F_{q\bar{q}g,T}^{D(3)}$ [all functions of $(x_{\mathbb{P}}, Q^2, \beta)$] are obtained by dividing out a factor $\frac{4\pi^2\alpha_{\text{em}}}{Q^2\beta}$. The first two terms in (4.1) are given in Eq. (2.18), they contain the LO impact factors with only a $q\bar{q}$ Fock component in the final state. The last term is the large Q^2 part of the contribution of the NLO impact factor: only the transverse part contributes, the longitudinal part being of higher twist. At large Q^2 the $q\bar{q}g$ -Fock state appears in a configuration in which the inter-quark-distance is tiny compared to $1/Q_s$ and one may take the corresponding coincidence limit. The $q\bar{q}$ part is then indistinguishable from a gluon and the analytic expression may be cast in terms of gluon dipole amplitudes $N_Y^{gg}(\mathbf{r}, \mathbf{b})$. The corresponding expression was first given by Golec-Biernat and Wüsthoff [31]. In it, masses are set to zero, in the spirit of a large Q^2 expansion. As with the first two terms of Eq. (4.1), which were already given in Eq. (2.18), we present this result in a notation inspired by [53], but with the b -integral not yet performed. This allows us to assess the quantitative impact of our lack of precise knowledge of the b -dependence in all three contributions. We retain the assumption that the dipole amplitudes only depend on $|\mathbf{r}|$ and are independent of the orientation of the dipole. Then

$$\begin{aligned} \frac{d\sigma_{q\bar{q}g,T}^{\gamma^* A \rightarrow Xp}}{d\beta} \Big|_{\text{LL}(Q^2)} &= \frac{\alpha_{\text{em}}\alpha_s C_f N_c}{8\pi^2 Q^2} \int_{\beta}^1 \frac{dz}{(1-z)^3} \left[\left(1 - \frac{\beta}{z}\right)^2 + \left(\frac{\beta}{z}\right)^2 \right] \int_0^{(1-z)Q^2} d\mathbf{k}^2 \ln \left(\frac{(1-z)Q^2}{\mathbf{k}^2} \right) \\ &\quad \times \int_0^{\infty} d\mathbf{r}^2 d\mathbf{r}'^2 \phi_{gg}(z, |\mathbf{k}|, |\mathbf{r}|, |\mathbf{r}'|) \int d^2\mathbf{b} N_Y^{gg}(\mathbf{r}, \mathbf{b}) N_Y^{gg}(\mathbf{r}', \mathbf{b}) \end{aligned} \quad (4.2a)$$

where

$$\phi_{gg}(z, |\mathbf{k}|, |\mathbf{r}|, |\mathbf{r}'|) := \mathbf{k}^4 J_2(|\mathbf{k}||\mathbf{r}|) K_2 \left(\sqrt{\frac{z}{1-z}} \mathbf{k}^2 \mathbf{r}^2 \right) K_2 \left(\sqrt{\frac{z}{1-z}} \mathbf{k}^2 \mathbf{r}'^2 \right) J_2(|\mathbf{k}||\mathbf{r}'|) . \quad (4.2b)$$

Any inclusion of quark masses (or a dependence of the dipole orientation) in such an expression would amount to a resummation of subleading effects with little control over their relevance. Similarly, we have no reliable argument to set the scale in the strong coupling α_s that appears in the prefactor of this expression. Both these assessments are reinforced once one starts analyzing the nonperturbative uncertainties in Eq. (4.1), even after it has been updated on the perturbative level with full NLO ingredients. To be consistent in our treatment below we will therefore also only consider the massless limit for the $q\bar{q}$ contributions.

In fact, the largest uncertainty in (4.2) and in the corresponding expressions for the quarks, Eq. (2.18), is related to the impact parameter integral which is not under perturbative control. Even after using data to set the overall normalization of the total cross section (the main non-perturbative parameter entering the LO total cross section), already the leading order diffractive contributions (the $q\bar{q}$ -terms in (4.1)) require additional nonperturbative input. This only gets more pronounced at NLO: higher order Fock components in the projectile wave function couple to higher n -point functions of Wilson lines, each of which is affected in its own way by non-perturbative

effects. This affects the relative normalizations of the terms in Eq. (4.1) as well as the relative normalization of total and diffractive cross sections. In practice, this manifests itself in a strong model dependence of the normalization of individual cross sections. The main issues here are the relative normalization of total and diffractive cross sections on the one hand and the weight of individual Fock components such as the $q\bar{q}$ and $q\bar{q}g$ contributions in Eq. (4.1) on the other.

4.2 Nonperturbative aspects of b -dependence and profile functions

The general attitude for the total cross section – to assume a fixed, x -independent target size, which at LO implies taking the $q\bar{q}$ -dipole amplitude as the product of a profile function $T(\mathbf{b})$ and r -dependent remainder $N_Y^{q\bar{q}}(\mathbf{r})$ (normalized to one at $r \rightarrow \infty$) leaves us with only a single fit parameter, the area resulting from the b -integration. In the diffractive case, the choice of profile strongly affects overall and relative normalization of

$$\int d^2\mathbf{b} N_Y^{q\bar{q}}(\mathbf{r}, \mathbf{b}) N_Y^{q\bar{q}}(\mathbf{r}', \mathbf{b}) \quad (4.3)$$

and

$$\int d^2\mathbf{b} N_Y^{gg}(\mathbf{r}, \mathbf{b}) N_Y^{gg}(\mathbf{r}', \mathbf{b}) \quad (4.4)$$

featuring in the formulae above as well as the more general correlators in their full NLO generalizations.

The simplest treatment would associate a factorized profile with each of the amplitudes, *identical* for both quarks and gluons according to (again with \mathcal{R} labeling the representation)

$$N_Y^{\mathcal{R}}(\mathbf{r}, \mathbf{b}) \rightarrow T(\mathbf{b}) N_Y^{\mathcal{R}}(\mathbf{r}) \quad (4.5)$$

and already in this case, the relative weight of diffractive and total cross sections are highly model dependent: A box profile $T_{\text{box}}(\mathbf{b})$ of height one, normalized in width to produce a factor

$$\sigma_0 = 2 \int d^2\mathbf{b} T(\mathbf{b}) \quad (4.6)$$

for the total cross section, results in a factor $\int d^2\mathbf{b} T_{\text{box}}^2(\mathbf{b}) = \sigma_0/2$ in the diffractive case. A Gaussian profile, which has some phenomenological justification at large $|\mathbf{b}|$,¹³ produces an additional factor $\frac{1}{2}$ in the diffractive case compared to the box profile: the area under $T_{\text{Gauss}}^2(\mathbf{b})$ is half the area under $T_{\text{Gauss}}(\mathbf{b})$. Clearly, the relative normalization of the total and the diffractive contributions is strongly dependent on the shape of the profile: Arbitrary factors of this sort can already be obtained by varying the width and the height of the box profile while keeping (4.6) fixed. While one may dismiss box profiles as unphysical, the issue remains: there is by no means a canonically prescribed physical profile that would outright eliminate such modeling choices.

Also a Gaussian profile, justified as it may be in some b -ranges for quarks, leaves intrinsic uncertainties: Not only is the overall normalization an issue, but also the relative normalization of parton

¹³This is based on successful parametrizations of meson production data via $e^{-B_d|t|} \left. \frac{d\sigma}{dt} \right|_{t=0}$ and constrains the shape for $|\mathbf{b}| > .3 \text{ fm}$ [72].

species (here quarks vs gluons) is affected. To explore this in more detail, we will invoke Casimir scaling, which we expect to hold at least at small b , as a guiding principle. Assuming the same profile factor in front of both quark and gluon amplitudes breaks Casimir scaling for the full b - and r -dependent amplitude $S_Y^{\mathcal{R}}(\mathbf{r}, \mathbf{b}) = 1 - N_Y^{\mathcal{R}}(\mathbf{r}, \mathbf{b})$ for effectively all $\mathbf{b}^2 > 0$ and hence, is not compatible with the notion that the energy dependence is dominated by perturbative gluon emission as encoded in the evolution equations and their Gaussian truncation. To guarantee this then requires that Casimir scaling is at best weakly broken for central collisions. Models with Casimir scaling restored completely are readily constructed: One may, for example, exponentiate the profile in the spirit of the IPSat and bCGC models according to

$$N_Y^{\mathcal{R}}(\mathbf{r}, \mathbf{b}) \rightarrow 1 - e^{-C_{\mathcal{R}} T(\mathbf{b}) \mathcal{G}_Y(\mathbf{r})} . \quad (4.7)$$

Alternatively, one might start with a factorized (Gaussian) profile for quarks to define $\mathcal{G}_Y(\mathbf{r}, \mathbf{b})$ for general use in the Gaussian truncation via $T(\mathbf{b}) N_Y^{q\bar{q}}(\mathbf{r}) =: 1 - e^{-C_f \mathcal{G}_Y(\mathbf{r}, \mathbf{b})}$. This allows to calculate general n -point functions in the Gaussian truncation with b -dependence and yields Casimir scaling for dipoles according to

$$N_Y^{\mathcal{R}}(\mathbf{r}, \mathbf{b}) \rightarrow 1 - [1 - T(\mathbf{b}) N_Y^{q\bar{q}}(\mathbf{r})]^{\frac{C_{\mathcal{R}}}{C_f}} . \quad (4.8)$$

Taking Casimir scaling as a guiding principle modifies both \mathbf{b} and \mathbf{r} dependence in a $C_{\mathcal{R}}$ dependent manner. In particular it leads to sizable changes in relative normalization of quark and gluon contributions as compared to the height-one box profile, for which all of these definitions are equivalent.

Plain exponentiation of the profile as in (4.7) enhances the influence from nonperturbative regions of phase space as Y increases: with increasing Y the large $|\mathbf{r}|$ -growth of $\mathcal{G}_Y(\mathbf{r})$ will progressively lift up any nonvanishing large $|\mathbf{b}|$ tails of the profile function if they exist at all. This in turn leads to Y -dependent growth of the overall normalization of the dipole cross section (after the b -integration is done). This results in an inconsistent, unphysical interplay of perturbative gluon emission with non-perturbative long range physics. One might attempt to regulate the large r -behavior of $\mathcal{G}_Y(\mathbf{r})$ to preclude that, but to do this in a defensible way would clearly require non-perturbative input completely outside the scope of JIMWLK evolution.

The ansatz (4.8) on the other hand does not require any such additional input: large $|\mathbf{r}|$ and $|\mathbf{b}|$ behavior decouple. Therefore we use this model below to estimate the impact on relative normalizations of cross sections.

With an eye to the total cross section, we first note that the model leads to representation dependent normalization for the b -integrated dipole cross section. Assuming a Gaussian profile, we find that the b -integrated profiles acquire a representation-dependent \mathbf{r} -dependence we denote $N_Y'^{\mathcal{R}}(\mathbf{r})$ as well as a nontrivial large $|\mathbf{r}|$ -normalization which we choose to display explicitly

$$\int d^2\mathbf{b} N_Y^{\mathcal{R}}(\mathbf{r}, \mathbf{b}) := 2\pi B_d H\left(\frac{C_{\mathcal{R}}}{C_f}\right) N_Y'^{\mathcal{R}}(\mathbf{r}) \xrightarrow{|\mathbf{r}| \rightarrow \infty} 2\pi B_d H\left(\frac{C_{\mathcal{R}}}{C_f}\right) \approx 2\pi B_d \begin{cases} 1 & \text{for quarks} \\ 1.6 & \text{for gluons} \end{cases} \quad (4.9)$$

where $H(z) := \psi_0(1 - z) + \gamma_E$ is the harmonic number of z (expressed via the digamma function ψ_0). Diffractive normalizations differ from the normalizations in the total cross section, and in addition, the \mathbf{r} - and \mathbf{r}' -dependence after b -integration for a generic representation only factorizes

approximately¹⁴

$$\int d^2\mathbf{b} N_Y^{\mathcal{R}}(\mathbf{r}, \mathbf{b}) N_Y^{\mathcal{R}}(\mathbf{r}', \mathbf{b}) \approx 2\pi B_d \left(2H\left(\frac{C_{\mathcal{R}}}{C_f}\right) - H\left(\frac{2C_{\mathcal{R}}}{C_f}\right) \right) \tilde{N}_Y^{\mathcal{R}}(\mathbf{r}) \tilde{N}_Y^{\mathcal{R}}(\mathbf{r}') \quad (4.10a)$$

with

$$\tilde{N}_Y^{\mathcal{R}}(\mathbf{r}) := \left[\frac{\int d^2\mathbf{b} [N_Y^{\mathcal{R}}(\mathbf{r}, \mathbf{b})]^2}{2\pi B_d \left(2H\left(\frac{C_{\mathcal{R}}}{C_f}\right) - H\left(\frac{2C_{\mathcal{R}}}{C_f}\right) \right)} \right]^{\frac{1}{2}}. \quad (4.10b)$$

Unsurprisingly, normalizations of quark and gluon contributions differ again: the relative diffractive normalizations for quarks and gluons are reliably assessed along the diagonal $\mathbf{r} = \mathbf{r}'$, where one finds

$$\int d^2\mathbf{b} [N_Y^{\mathcal{R}}(\mathbf{r}, \mathbf{b})]^2 = 2\pi B_d \left(2H\left(\frac{C_{\mathcal{R}}}{C_f}\right) - H\left(\frac{2C_{\mathcal{R}}}{C_f}\right) \right) \xrightarrow{|\mathbf{r}| \rightarrow \infty} 2\pi B_d \begin{cases} \frac{1}{2} & \text{for quarks} \\ 1. & \text{for gluons} \end{cases}, \quad (4.11)$$

i.e. the gluon contribution is enhanced by a factor of two compared the quarks. This is not the case for the ansatz (4.5) or the use of a box profile of height one. We take this as a practical indication that the nonperturbative uncertainties are indeed large and that a precision fit at NLO requires refined nonperturbative input or independent phenomenological constraints on the individual normalization of each Fock component.

On the practical side one finds that factorization is exact for quarks and a good approximation for gluons. For quarks $\tilde{N}_Y^{q\bar{q}}(\mathbf{r}) = N_Y^{q\bar{q}}(\mathbf{r})$ and $(2H(\frac{C_f}{C_f}) - H(\frac{2C_f}{C_f})) = \frac{1}{2}$. For gluons one may, for simplicity, even use $N_Y'^{gg}(\mathbf{r})$ as inspired by (4.9) to approximate

$$\int d^2\mathbf{b} N_Y^{gg}(\mathbf{r}, \mathbf{b}) N_Y^{gg}(\mathbf{r}', \mathbf{b}) \approx 2\pi B_d C N_Y'^{gg}(\mathbf{r}) N_Y'^{gg}(\mathbf{r}') \quad (4.12)$$

where $C = 1$, in line with (4.11).

These may be used in the momentum space formulae originally derived with height one box profiles (which factorize trivially) after properly including the normalization factors – the factorization error is much smaller than any nonperturbative uncertainty inherent in the choice of the profile model.

4.3 Diffractive fits to HERA data

Given the uncertainties arising both from incomplete NLO impact factors (Sec. 4.1) and nonperturbative aspects of the impact parameter dependence (Sec. 4.2) precision fit to diffractive data are out of the question.

The fits presented below are done with this in mind – mainly to assess if one can get a qualitative agreement with data, based on the fit parameters Y_{off} and Λ obtained from the total cross section. To optimize the diffractive fits, an independent normalization B_d is allowed and its relation to $B_d^{\text{tot}} =$

¹⁴For (4.7) factorization of \mathbf{r} -dependence is generically a bad approximation and the result depends strongly on the IR regularization at large r .

$\sigma_0/4\pi$ is ignored in the fit. How closely the fit results match is treated as a consistency check. As already indicated, the scale of α_s in the $q\bar{q}g$ component of the diffractive cross section in Eq. (4.2a) is undetermined. This factor of α_s will therefore be treated as a further fit parameter. While B_d and α_s must be meaningful in the context of HERA physics, we need a means to accommodate the uncertainties exposed above. In this sense, we take the precise values of these parameters as scenario dependent.

In the fits to the total cross section, the differences between the different scenarios using the energy conservation correction turned out to be small. For this reason, it is sufficient to consider only GT results from our simulations. We retain the GB-W model results for comparison. In what follows, the overall normalization of the $q\bar{q}$ components is the same in all cases since for the quarks the profile function factorizes trivially. For the gluon component, the relative normalization is scenario dependent and we explicitly consider the following two cases:

- **[fact]**: b -dependence completely factorized: the simplest case given in (4.5). The $q\bar{q}$ and $q\bar{q}g$ components are equally weighted.
- **[sc]**: b -dependence based on Casimir scaling after using the approximation of Eq. (4.10). The $q\bar{q}g$ component is enhanced by a factor of two relative to the ansatz (4.5), its shape is slightly modified. (The $q\bar{q}$ components remain completely factorized.)

$N_Y^{gg}(\mathbf{r})$ for the two schemes and the corresponding momentum space amplitudes $\mathcal{N}_Y^{gg}(\mathbf{q})$ (see App. D.1 for definitions), are illustrated in Fig. 12. The difference between the factorized and

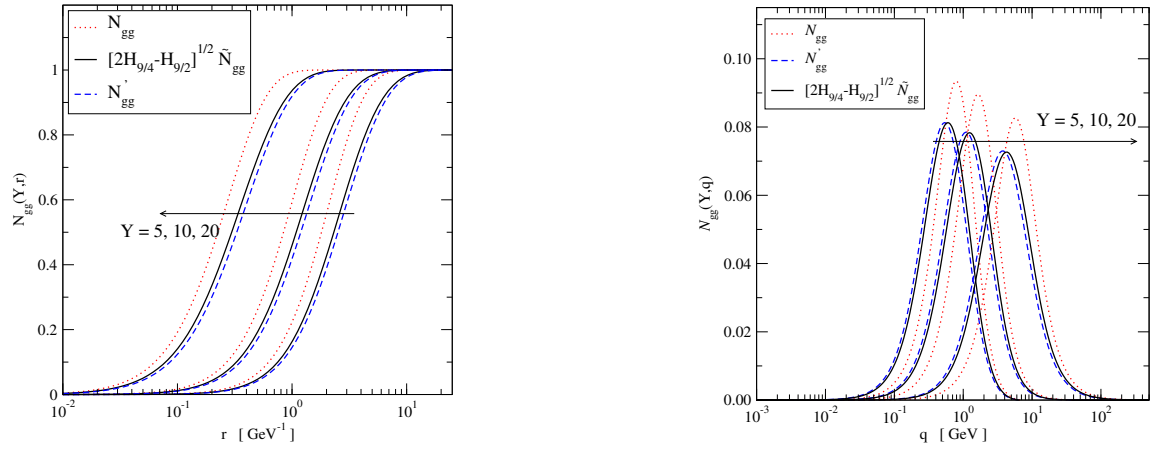


Fig. 12: Three gluon dipole inputs: globally factorized $N_Y^{gg}(\mathbf{r})$ Eq. (4.5), Casimir scaling based b -integrated $(2H_{9/4} - H_{9/2})^{1/2} \tilde{N}_Y^{gg}(\mathbf{r})$ Eq. (4.10) and $N_Y^{gg}(\mathbf{r})$ in coordinate space (**left panel**) and momentum space (see App. D.2 for precise definitions) (**right panel**).

Casimir scaling schemes are pronounced: The main effect of [sc] is to increase the effective R_s after b -integration is done compared to the factorized scheme [fact].¹⁵ As a result the [sc]-gluon stays closer to the quark result than the [fact]-gluon, see Fig. 24 in App. D.

¹⁵We note in passing that the gluon amplitudes $N_Y^{gg}(\mathbf{r})$ (dashed blue) and $\tilde{N}_Y^{gg}(\mathbf{r})$ (solid black) are almost the same: a tiny rescaling in r is required in order to (perfectly) match the curves.

Since the literature lists data directly for structure functions $x_{\mathbb{P}} F_2^{D(3)} = \frac{Q^2 \beta}{(2\pi)^2 \alpha_{\text{em}}} \frac{d\sigma^{\gamma^* A \rightarrow X_{\mathbb{P}}}}{d\beta}$, the actual task is to compute each component of the cross section Eq. (4.1) with the parameters obtained from the total cross section and then to optimize B_d and α_s in the following expression,

$$x_{\mathbb{P}} F_2^{D(3)} = B_d \left[x_{\mathbb{P}} F_{q\bar{q},T}^{D(3)} \big|_{B_d=1} + x_{\mathbb{P}} F_{q\bar{q},L}^{D(3)} \big|_{B_d=1} + \alpha_s x_{\mathbb{P}} F_{q\bar{q}g,T}^{D(3)} \big|_{B_d=1, \alpha_s=1} \right]. \quad (4.13)$$

The numerical burden of this procedure can be vastly reduced by replacing the coordinate space expressions for these contributions shown in the text by their momentum space counterparts, see App. D. $x_{\mathbb{P}} F_2^{D(3)}$ and its components are visualized in Fig. 13: the overall behavior of $x_{\mathbb{P}} F_2^{D(3)}$ is analogous in all cases, including the GB-W model.

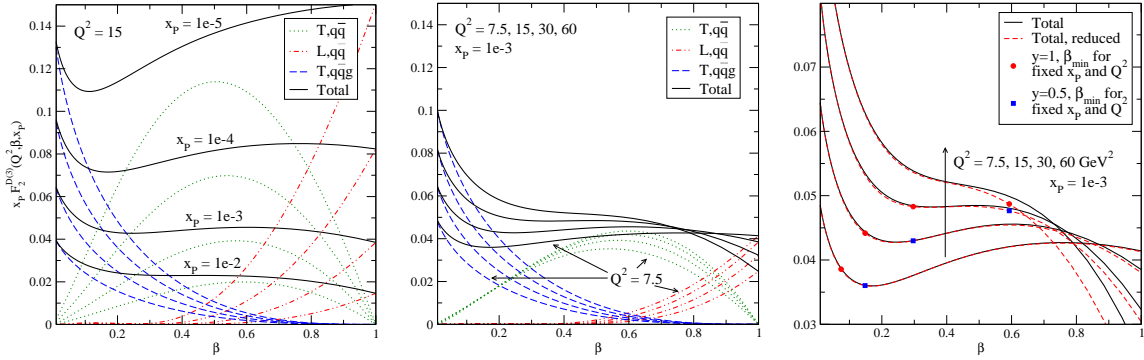


Fig. 13: Behavior of the contributions given in Eq. (D.2). **Left:** Change in $x_{\mathbb{P}} F_2^{D(3)}$ as $x_{\mathbb{P}}$ is decreased with fixed Q^2 . **Middle:** Q^2 is increased with fixed $x_{\mathbb{P}}$. **Right:** A comparison of $x_{\mathbb{P}} \sigma_r^{D(3)}$ and $x_{\mathbb{P}} F_2^{D(3)}$ according to Eq. 4.14. See the text for more details.

$N_Y^{\mathcal{R}}(\mathbf{r})$: (4.5)	GB-W	GT	$\tilde{N}_Y^{\mathcal{R}}(\mathbf{r})$: (4.10)	GT
B_d	2.85	4.70	B_d	4.70
α_s	0.54	0.44	α_s	0.31
$\chi^2/552$	1.44	1.44	$\chi^2/552$	1.47

Table 3: The results of the fits to the diffractive data [37–41]. **Left:** [fact] Fully factorized b -dependence of Eq. (4.5). **Right:** [sc] b -dependence based on Casimir scaling according to Eqs. (4.10), (4.11) and (4.12).

Fig. 14 presents a subset¹⁶ of $x_{\mathbb{P}} F_2^{D(3)}(Q^2, \beta, x_{\mathbb{P}})$ data from H1 [39] together with the theoretical predictions. We show three practically indistinguishable scenarios:

- [fact]: GT with the factorized ansatz (4.5) (denoted “gt, fact.”)
- [sc]: GT based on the Casimir scaling ansatz of (4.10) (denoted “gt, sc.”).
- GB-W with the factorized ansatz (4.5) (denoted “gbw, fact.”)

¹⁶Although not shown, the data from ZEUS [37, 38, 40, 41] were also included in the fits.

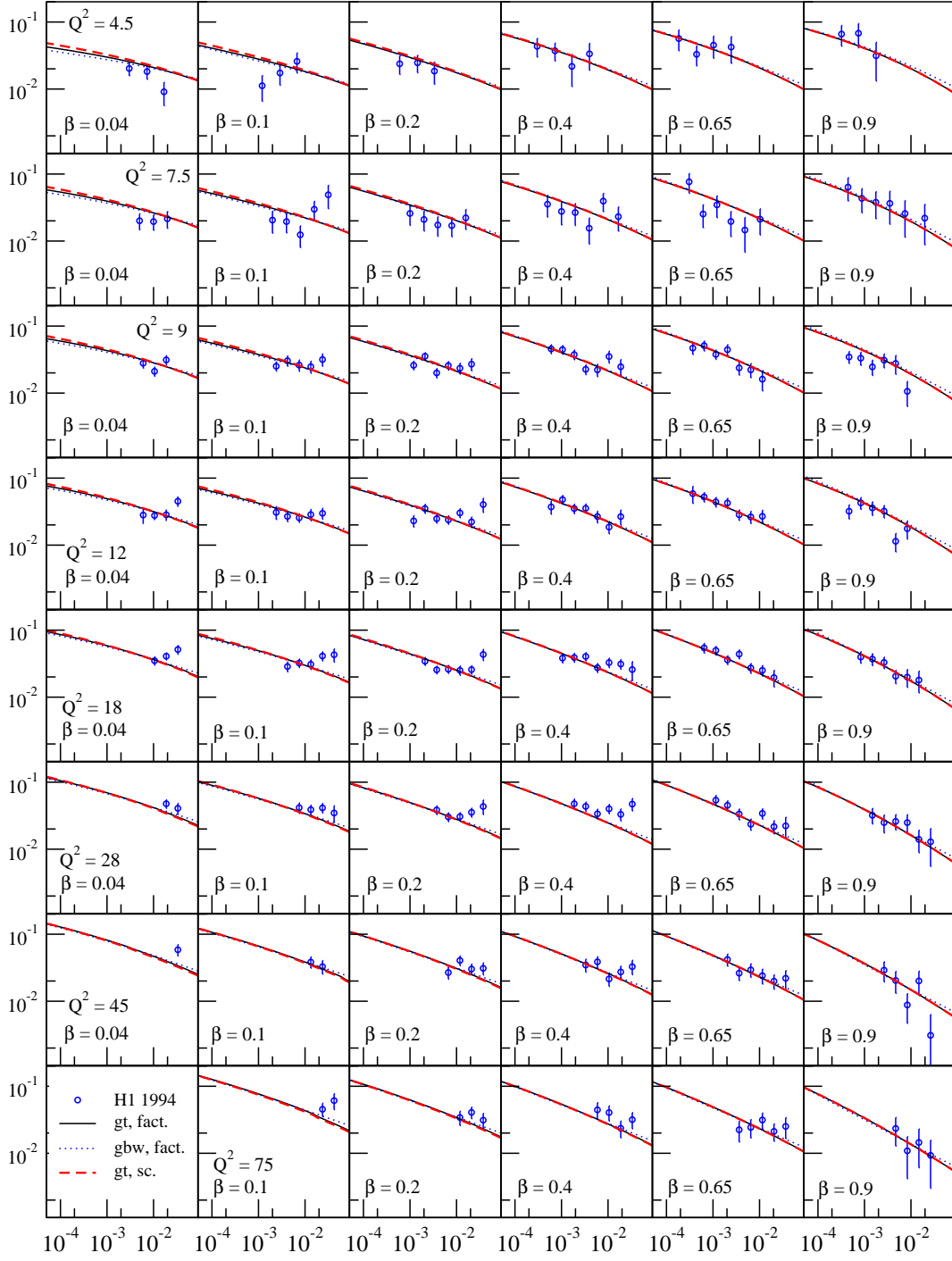


Fig. 14: $x_{\mathbb{P}} F_2^{D(3)}(x_{\mathbb{P}} Q^2, \beta)$ as a function of $x_{\mathbb{P}}$. The data are from H1 [39].

GT	$\beta \leq 0.5$	$\beta > 0.5$
$N_Y^{\mathcal{R}}(\mathbf{r})$: (4.5)	$\chi^2/264 = 1.58$	$\chi^2/288 = 1.30$
$\tilde{N}_Y^{\mathcal{R}}(\mathbf{r})$: (4.10)	$\chi^2/264 = 1.64$	$\chi^2/288 = 1.31$

Table 4: The diffractive data split into two subsets $\beta \leq 0.5$ and $\beta > 0.5$. The data with $\beta > 0.5$ fit somewhat better (a region where $q\bar{q}$ components dominate).

The corresponding fit parameters are presented in Tables 3.

The parameter α_s turns out to be strongly scenario dependent, reflecting the differences in the shapes, scales and overall normalizations of the gluon amplitudes shown in Fig. 12. For GT, the common normalization B_d as well as the fit quality χ^2/dof are of the same order in all cases. The values of B_d are systematically smaller than the experimental value $B_d = 7.1 \text{ GeV}^{-2}$ reported in [38] but, they *are* consistent with $B_d^{\text{tot}} = \sigma_0/4\pi$ obtained from the total cross section.¹⁷ BK evolution (not shown) results in fits of the same systematic behavior and fit quality as the GT fits presented here.

Fit quality varies noticeably across phase space: splitting the data into two subsets $\beta \leq 0.5$ and $\beta > 0.5$ or equivalently $Q^2 \leq M_X^2$ and $Q^2 > M_X^2$,¹⁸ we observe that the best match with the data is obtained at large β where the $q\bar{q}$ components dominate, see Table 4. This reinforces our statements that a better treatment of the $q\bar{q}g$ amplitude is required. The presently implemented improvements in our treatment of the b -dependence only lead to a tiny improvements in fit quality: the main differences between the gluon amplitudes resulting from the two schemes dubbed [fact] and [sc] are effectively absorbed into the normalization of the $q\bar{q}g$ component via parameter α_s .

As a check of consistency, the fits are also performed by using the reduced cross section (as was done in [53])

$$\sigma_r^{D(3)} = F_2^{D(3)} \left(1 - \frac{y^2}{1 + (1-y)^2} \frac{F_{q\bar{q},L}^{D(3)}}{F_2^{D(3)}} \right) \quad ; \quad y \approx \frac{Q^2}{s_{ep}x_{\text{bj}}} \quad , \quad (4.14)$$

where $s_{ep} = 318^2 \text{ GeV}^2$ at HERA. We note that for the present diffractive data $y < 0.5$ so that $\sigma_r^{D(3)}$ hardly deviates from $F_2^{D(3)}$ over the bulk of the data range¹⁹. As a consequence, one would at best expect a tiny improvement in the fit quality in any approach once the reduced cross section is used. We find indeed that χ^2 values improve almost imperceptibly to $\chi^2/552 = 1.39$ and $\chi^2/552 = 1.42$, respectively. Consistent with this, the parameters remain practically unchanged: $B_d = 4.75 \text{ GeV}^{-2}$ and $\alpha_s = 0.43$ for (4.5) and $B_d = 4.75 \text{ GeV}^{-2}$ and $\alpha_s = 0.31$ for (4.10).

To pinpoint precisely where the differences arise consider the rightmost panel in Fig. 13: this compares $\sigma_r^{D(3)}$ with $F_2^{D(3)}$ for fixed $x_{\mathbb{P}}$ and Q^2 as a function of β . It should be noted that if

¹⁷Note that the use of $T_{\text{box}}(\mathbf{b})$ of height one induces an overall factor of two in *all* components relative to $T_{\text{Gauss}}(\mathbf{b})$ in the framework of (4.5). Since it can only be absorbed by a redefinition of B_d in the diffractive fits, this would lead to major inconsistencies.

¹⁸As with the data for the total cross section, there is kinematical correlation in the data range: β is small if $Q^2 \ll M_X^2$. In this case, $x_{\text{bj}} = x_{\mathbb{P}}\beta$ is also small and thus $s_{\gamma^*p} \approx Q^2/x_{\text{bj}}$ is large. Since $s_{\gamma^*p} \ll s_{ep}$ in the experiments is finite, small values of β are more likely paired with low Q^2 .

¹⁹For 322 points out of 512 $y < 1/4$ implying that the factor in front of $F_{q\bar{q},L}^{D(3)}$ in (4.14) is < 0.04 . The data with $1/4 \leq y \leq 1/2$ are most likely associated with large Q^2 .

the diffractive phase space is presented in this way, one must take into account that the kinematic limit $y \leq 1$ is violated at small β . Two conditions $y = 1$ and $y = 0.5$ marked by the red circles and blue squares overlaid on each curve are indicating the smallest possible β for fixed $x_{\mathbb{P}} = 10^{-3}$, $Q^2 = \{7.5, 15, 30, 60\}$ GeV² and $s_{ep} = 318^2$ GeV² (set by $\beta_{\min} = Q^2/(x_{\mathbb{P}}s_{ep}y)$). The former is the absolute kinematic limit of HERA whereas the latter approximately marks the lowest values of β of the existing data.

4.4 Ratios with the total cross section

The approximate constant ratio of the diffractive to inclusive cross sections was originally observed at HERA by ZEUS collaboration [40]. To compute this observable, we use the relation (see [40, 41])

$$\frac{1}{2M_X} \frac{d\sigma_{\gamma^*p \rightarrow Xp}^{\text{diff}}(Q^2, M_X, x_{\mathbb{P}})}{dM_X} \approx \frac{(2\pi)^2 \alpha_{\text{em}}}{Q^2(Q^2 + M_X^2)} x_{\mathbb{P}} F_2^{D(3)}(Q^2, M_X, x_{\mathbb{P}}) \quad (4.15)$$

for the diffractive cross section. The ratio of the diffractive cross section to the total cross section can then be computed by

$$R_{\text{tot}}^{\text{diff}}(s_{\gamma^*p}) = \frac{\int_{M_a}^{M_b} dM_X d\sigma_{\gamma^*p \rightarrow Xp}^{\text{diff}}(Q^2, M_X, x_{\mathbb{P}})/dM_X}{\sigma_{\gamma^*p}^{\text{tot}}(x_{\text{bj}}, Q^2)} \quad ; \quad Q^2 \ll s_{\gamma^*p} \ll s_{ep} \quad , \quad (4.16)$$

where the integration boundaries, i.e. the bins for the diffractive mass M_X , are determined by the experimental setup (see Table. 5).

In the calculation of $R_{\text{tot}}^{\text{diff}}$ given above, no new free parameters are introduced. Basically, $R_{\text{tot}}^{\text{diff}}$ is parameter free since the normalizations cancel trivially. However, as mentioned earlier, we ignored the value of $B_d^{\text{tot}} = \sigma_0/4\pi$ from the total cross section and allowed a distinct normalization in the diffractive fits. For instance, in the case of GT the ratio of the optimal normalizations is found to be $B_d/B_d^{\text{tot}} \approx 1.07$, i.e. fairly close to one. In practice this means that *all* HERA data can be resolved by the same normalization $B_d^{\text{tot}} = \sigma_0/4\pi \approx 4.40$ GeV⁻²: a single parameter fit to the diffractive cross section then results in $\alpha_s = 0.34$ with a reasonable $\chi^2/552 = 1.52$ for the scenario (4.10). The region $\beta \leq 0.5$ is responsible of the reduction of the overall χ^2/dof : at $\beta > 0.5$ the fit quality is actually slightly improved.

In Fig. 15 are shown the latest $R_{\text{tot}}^{\text{diff}}$ data from [41] together with the theoretical predictions based on the fits to the inclusive and diffractive cross sections presented earlier²⁰. As seen, the overall behavior of $R_{\text{tot}}^{\text{diff}}$ is well produced with fairly good fit qualities (see Table 5). In the case of the reduced cross section $\sigma_r^{D(3)}$ only the bins $M_X \in [0.28, 2]$ and $[2, 4]$ GeV (the largest β) are affected, resulting in a slightly better χ^2/dof . In Fig. 16 we show the corresponding total cross section alone: the match with the data is strikingly good indicating that the source of the uncertainties in $R_{\text{tot}}^{\text{diff}}$ is the incomplete description for the diffractive cross section.

²⁰More $R_{\text{tot}}^{\text{diff}}$ data with $Q^2 \in [2.7, 55]$ GeV² can be found in [73]. The fits to this data turned out to be slightly worse.

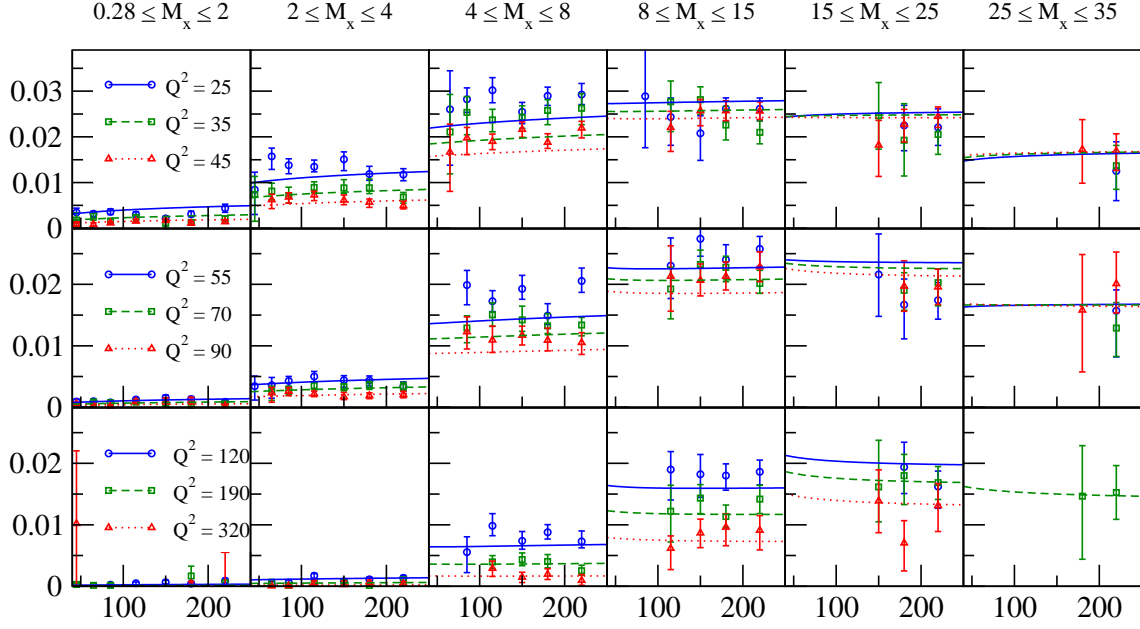


Fig. 15: The ratio of the diffractive versus the inclusive cross sections according to Eq. 4.16 as a function of $\sqrt{s_{\gamma^*p}}$ for different values of Q^2 and bins of diffractive mass M_X . The smallest values of β are located in the top right corner: from there, following the panels to the left or down increases β . The ratios are based on the parameters for GT in Tables 1 (σ_{tot}) and 3 (σ_{diff} , Eq. (4.10)). The data are from [41].

M_X [GeV]	0.28-2	2-4	4-8	8-15	15-25	25-35	Σ	
Treatment	dof	54	57	46	37	23	10	227
$F_2^{D(3)}$; $N_Y^{\mathcal{R}}(\mathbf{r})$: (4.5)		2.05	1.22	2.27	0.77	1.16	0.29	1.51
$F_2^{D(3)}$; $\tilde{N}_Y^{\mathcal{R}}(\mathbf{r})$: (4.10)		2.05	1.22	2.30	0.82	0.88	0.24	1.49
$\sigma_r^{D(3)}$; $N_Y^{\mathcal{R}}(\mathbf{r})$: (4.5)		1.78	1.16	2.30	0.77	1.16	0.29	1.44
$\sigma_r^{D(3)}$; $\tilde{N}_Y^{\mathcal{R}}(\mathbf{r})$: (4.10)		1.78	1.16	2.33	0.82	0.88	0.23	1.42

Table 5: The fit results of $R_{\text{tot}}^{\text{diff}}$ for each M_x bin (GT, diffractive cross section vs. reduced cross section). The data are from [41].

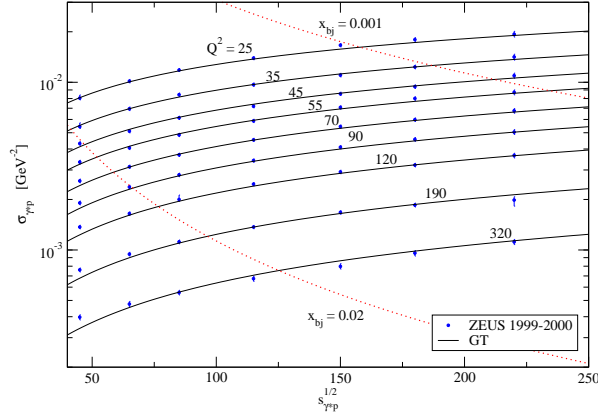


Fig. 16: The γ^*p cross section as a function of $\sqrt{s_{\gamma^*p}} = \sqrt{Q^2(1/x_{bj} - 1)}$ for different values of Q^2 . The data are from [41], i.e. it is the total cross section part of $R_{\text{tot}}^{\text{diff}}$ data shown in Fig. 15. As long as $x_{bj} = Q^2/(Q^2 + s_{\gamma^*p}) \lesssim 0.02$ the theoretical predictions match excellently with the experiments. The breakdown at large x_{bj} becomes evident as one approaches the lower left corner with fixed $\sqrt{s_{\gamma^*p}} \approx 45 - 65$ GeV.

5 Asymptotic versus pre-asymptotic fits, a comparison

As we have discussed at length for the total cross section in Sec. 3, an asymptotic fit without the energy conservation included is not feasible. This is caused by a strong fit tension arising between too large an evolution speed at small Y and too shallow a dipole correlator shape at large Y . Once energy conservation is included, both features improve towards what is needed to match the data: evolution speed is uniformly lowered and correlators become steeper overall. The upshot is a fit whose main ingredients are determined perturbatively – both shape and evolution speed in the asymptotic region are predominantly²¹ determined by the nonlinear structure of the evolution equation and its kernel. As reported, it works flawlessly up to Q^2 of 1200 GeV², i.e. the largest Q^2 values available at HERA for $x \leq 0.02$.

The fits performed in [45] by contrast include running coupling effects but no energy conservation correction. They obtain a fit that is almost as good as the GT fits described in this paper. To achieve this, they *must* move away from the asymptotic pseudo-scaling region and use pre-asymptotic features of an evolution trajectory. This allows them to simultaneously satisfy the speed and steepness requirements of the dipole correlators and improve the fit quality over what is possible in the asymptotic (pseudo-scaling) domain without the energy conservation correction included. The features of the correlators along the part of the evolution trajectory used in the fit differ drastically from those in the asymptotic case and a detailed comparison is in order. We will perform this comparison for both the total and diffractive cross sections (which [45] does *not* consider), to illustrate once more that, with present theoretical limitations, there is no hope to use the more differential diffractive cross section to differentiate between theoretical approaches.

In doing so we have to deal with secondary differences of the fit procedures in our case and in [45]. Where we have argued for the use of current quark masses and have included a charm quark

²¹Aside from regulator effects on the coupling which are visible due to the R_s values inherent to the kinematic properties of the HERA experiments.

contribution successfully using (optionally) an m_f -independent remapping of $x \mapsto x_{\text{eff}} = x \frac{Q^2 + Q_s^2(x)}{Q^2}$, [45] uses three quarks with $m_f = 140$ MeV and an m_f -dependent remapping of $x \mapsto x_{\text{eff}}^{\text{GB-W}} = x \frac{Q^2 + 4m_f^2}{Q^2}$, which, if used with our solutions degrades the χ^2 considerably. The fit and parameters of [45] were done for $x \leq 0.01$ instead of the $x \leq 0.02$ we have used. For comparison we have to restrict ourselves to the same range. This reduces the Q^2 range of available data from $Q^2 \leq 1200 \text{ GeV}^2$ to $Q^2 \leq 650 \text{ GeV}^2$. Fortunately the separation scheme used to define the running coupling contribution is the same in both treatments.

For the pre-asymptotic fit scenario, inclusion of the charm quark is virtually impossible without reworking most of the ingredients. To simplify the comparison we have also left out the charm quark in the asymptotic fits we compare to explicitly – for its inclusion see Sec. 3.3.

We have attempted to expose the effects of the secondary differences in the fit procedures by comparing a set of fit scenarios that permute some of these ingredients. The fits below are tagged as $\text{BK}_{\text{phys}}^{\text{A}}$, $\text{BK}_{140}^{\text{A}}$, $\text{BK}_{140}^{\text{A},x_{\text{eff}}}$, $[\text{BK}_{140}^{\text{A},x_{\text{eff}}}]_{\text{rc}}$, and $[\text{BK}_{140}^{\text{P},x_{\text{eff}}}]_{\text{rc}}$ and defined as follows

- $\text{BK}_{\text{phys}}^{\text{A}}$: asymptotic (pseudo-scaling), energy conservation correction included, three quarks, mass pattern $m_f = \{3, 5, 105\}$ MeV. An x remapping induces no discernible differences.
- $\text{BK}_{140}^{\text{A}}$: asymptotic (pseudo-scaling), energy conservation correction included, three quarks, mass pattern $m_f = \{140, 140, 140\}$ MeV.
- $\text{BK}_{140}^{\text{A},x_{\text{eff}}}$: asymptotic (pseudo-scaling), energy conservation correction included, three quarks, mass pattern $m_f = \{140, 140, 140\}$ MeV, remapping of x according to $x \mapsto x_{\text{eff}}^{\text{GB-W}} = x \frac{Q^2 + 4m_f^2}{Q^2}$.
- $[\text{BK}_{140}^{\text{A},x_{\text{eff}}}]_{\text{rc}}$: same as previous, but with the energy conservation correction omitted. This attempts a fit on the asymptotic line the fit of [45] eventually merges onto.
- $[\text{BK}_{140}^{\text{P},x_{\text{eff}}}]_{\text{rc}}$, the fit developed in [45]: pre-asymptotic, NLO corrections restricted to running coupling corrections (energy conservation correction excluded), three quarks, mass pattern $m_f = \{140, 140, 140\}$ MeV, remapping of x according to $x \mapsto x_{\text{eff}}^{\text{GB-W}} = x \frac{Q^2 + 4m_f^2}{Q^2}$. Evolution starts from a GB-W like initial state $S_{Y_0=\ln(1/0.01)}^{q\bar{q}}(r) = \exp[-(rQ_{s,0}/2)^2]$ with $Q_{s,0}^2 = 0.241 \text{ GeV}^2$. The scale choice for the running coupling in the fit is parametrized differently from our treatment: In the argument of a running coupling we use an r dependent scale in the form $\frac{\mu^2(r)}{\Lambda^2} = \frac{C^2}{r^2\Lambda^2}$ with $C^2 = 4e^{-2\gamma_E - \frac{5}{3}}$ and vary Λ . They use $\frac{\mu^2(r)}{\Lambda^2} = \frac{4C^2}{r^2\Lambda^2}$, with Λ set to .241 GeV a priori, and obtain $C^2 = 5.3$ from the fit, see Fig. 17, right panel.

In all treatments including $[\text{BK}_{140}^{\text{P},x_{\text{eff}}}]_{\text{rc}}$ an APT regulator is used while [45] regulate the coupling by freezing it at $\alpha_s^{\text{max}} = .7$, see Fig. 17, right panel. Note that the choice of initial condition in $[\text{BK}_{140}^{\text{P},x_{\text{eff}}}]_{\text{rc}}$ is quite restrictive and no effort is made to vary the shape of the correlator other than allowing for some offset rapidity before matching evolution results to data. A systematic study of the impact of varying the shape, even a theoretical exploration of which shape features might be responsible for what kind of physics property of the cross section is still outstanding. Here we only attempt to contrast asymptotic fits with *one* example of a preasymptotic one.

$x_{bj} \leq 0.01$		BK_{phys}^A	BK_{140}^A	$BK_{140}^{A,x_{\text{eff}}}$	$[BK_{140}^{A,x_{\text{eff}}}]_{\text{rc}}$	$[BK_{140}^{P,x_{\text{eff}}}]_{\text{rc}}$
$Q^2 \leq 45$ GeV ²	Λ [MeV]	87.3	59.0	52.7	31.2	241
	$\chi^2/200$	0.82	0.95	1.01	2.23	0.97
	σ_0 [GeV ⁻²]	55.68	72.3	77.3	79.3	81.6
$Q^2 \leq 650$ GeV ²	Λ [MeV]	93.7	68.9	63.1	52.4	241
	$\chi^2/230$	0.86	1.02	1.09	3.42	1.01
	σ_0 [GeV ⁻²]	55.91	69.9	74.2	71.4	81.7

Table 6: A comparison of the fits to the total cross section. Parameters are based on the full range $x_{bj} \leq .01$, $Q^2 \leq 650$ MeV. The superscripts “A” and “P” refer to the asymptotic and pre-asymptotic fits, respectively. The quark masses are $m_{u,d,s} = \{3, 5, 105\}$ MeV (physical) or $m_{u,d,s} = \{140, 140, 140\}$ MeV (ad hoc).

We intentionally only show BK based fits to allow for a direct comparison, despite the fact that GT fits have better χ^2 and note that the modifications from BK_{phys}^A through BK_{140}^A to $[BK_{140}^{A,x_{\text{eff}}}]_{\text{rc}}$ incrementally reduce fit quality. Fit quality recovers only for $[BK_{140}^{P,x_{\text{eff}}}]_{\text{rc}}$ and relies on the freedom gained once one allows for correlator shapes away from the pseudo-scaling behavior. See Table 6. The first qualitative difference of the fit scenarios is captured in a plot of evolution speeds (left panel in Fig. 17) which indicates the parts of trajectories used in the fit of [45], $[BK_{140}^{P,x_{\text{eff}}}]_{\text{rc}}$, and our favored asymptotic fits. The $R_s\Lambda$ ranges of the successful pre-asymptotic and asymptotic fits are strongly shifted against each other as required by the constraints on evolution speeds. Note that the fit interval matched onto data on the pre-asymptotic fit trajectory ends before the

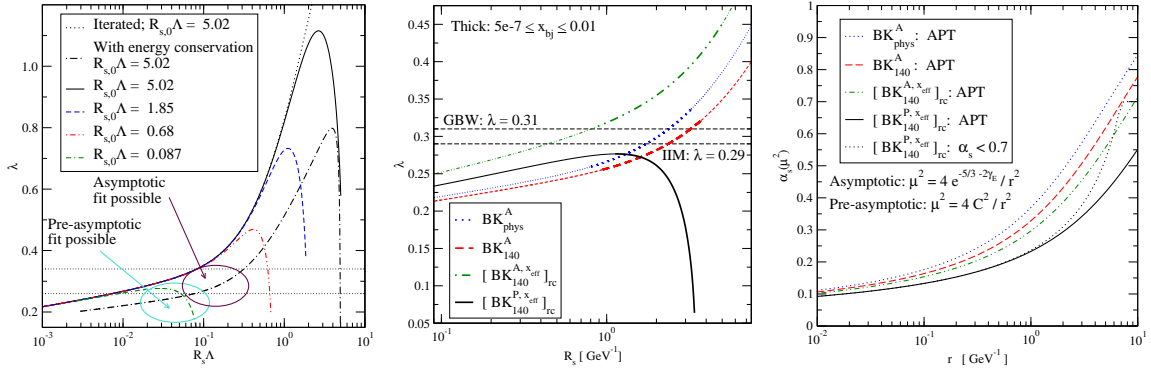


Fig. 17: Left: Comparison of evolution speeds of pre-asymptotic fits without energy conservation correction [45] with the asymptotic fits including the energy conservation correction. The marked areas indicate the ranges of $\lambda(R_s(Y)\Lambda)$ where the different fit strategies match to data. The curve starting from $R_{s,0}\Lambda = 0.087$ corresponds to the pre-asymptotic fit. **Middle:** The situation after the ratio C/Λ is known from the data fit. Due to the unique energy dependence of each scenario, the fit range $5 \times 10^{-7} \leq x_{bj} \leq 0.01$ is emphasized by thicker line width for each curve. **Right** $\alpha_s(\mu)$ for each scenario after the ratio C/Λ is determined by the data fit. $C/\Lambda \approx 5.2, 7.1$ and 19.1 for the scenarios BK_{phys}^A , BK_{140}^A and $[BK_{140}^{A,x_{\text{eff}}}]_{\text{rc}}$, respectively. This gives an impression of the size of the slowdown effect originating from the energy conservation correction. The black dashed line shows coupling used in [45], regulated in the IR by freezing it at maximum value of .7.

asymptotic shape is reached (and the correlator shape would have become too shallow): the fit

strategy applied in the pre-asymptotic fit $[\text{BK}_{140}^{\text{P},x_{\text{eff}}}]_{\text{rc}}$ [45] relies on features of the initial condition *throughout*. Strikingly, the evolution speed increases over the whole fit interval of the pre-asymptotic fit, while it monotonically decreases in the asymptotic case. In addition, the dipole cross section of $[\text{BK}_{140}^{\text{P},x_{\text{eff}}}]_{\text{rc}}$ interpolates between a steep, almost GB-W shape at $x_{\text{bj}} = .01$, and a final shape at the small x_{bj} end of the data range that remains still steeper than the (too shallow) pseudo-scaling shape obtained from running coupling BK evolution without the energy conservation correction. (See Fig. 18, left panel.) Note that detailed short distance information in this fit is predominantly

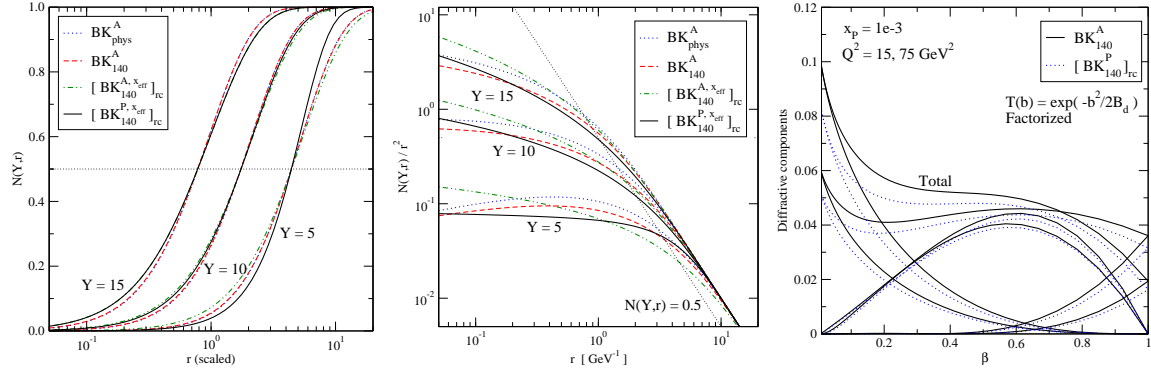


Fig. 18: Correlator shapes obtained from the fits to the total cross section (Table 6). **Left:** A comparison of the shapes of $N_Y^{q\bar{q}}(r)$. The curves are shifted to match at $N_Y^{q\bar{q}}(r) = 0.5$ for each value of Y . **Middle:** $N_Y^{q\bar{q}}(r)/r^2$: the behavior at small r . **Right:** Contributions to the diffractive cross section for $x_{\text{P}} = 10^{-3}$ using a factorized b -profile.

imprinted through the choice of initial condition. Evolution effects, which propagate into the UV just like in the linear BFKL case, have not yet reached the short distance tail of the correlators. This is different in our asymptotic fits, where the shape is determined entirely by the structure of r.h.s. of the evolution equation, which in its entirety is based on a (highly resummed) perturbative calculation.²² To highlight the actual differences at short distances we show part of the short distance asymptotics of the fits by plotting $\frac{N_Y^{q\bar{q}}(r)}{r^2}$ in Fig. 18, right panel.²³

We conclude our comparison with a look at the diffractive cross sections, see Table 7. The pattern

$x_{\text{P}} \leq 0.01$	B_d	α_s	$\beta \leq 0.5$	$\beta > 0.5$	Total	$Q^2 \leq 10$	$10 < Q^2 \leq 45$	$Q^2 > 45$
χ^2/dof			$\chi^2/153$	$\chi^2/215$	$\chi^2/368$	$\chi^2/97$	$\chi^2/174$	$\chi^2/97$
$\text{BK}_{140}^{\text{A}}$, fact.	6.30	0.40	1.39	1.37	1.38	1.53	1.35	1.48
$[\text{BK}_{140}^{\text{P}}]_{\text{rc}}$ fact.	7.47	0.34	1.29	1.22	1.25	1.50	1.17	1.15

Table 7: The fits to the diffractive data with the parameters shown in Table 6. Tag “fact.” refers to the globally factorized case (4.5), however, with the large- N_c replacement $C_f \rightarrow N_c/2$.

of fit quality shown in the table repeats what we have already seen for the total cross section:

²²Operationally this is imprinted on the solution by having the solution evolve into the asymptotic region over a long Y interval before the fit range is reached and the correlators have reached pseudo-scaling shapes. Practically this will only affect the UV within the BFKL diffusion radius, but this easily covers the range relevant to HERA fits.

²³For a direct comparison in momentum space, consult App. D.2, in particular Fig. 24, both the middle and right panels.

- The asymptotic fits shown in Sec. 3 and 4 show a good fit quality, also when restricted to $x_{\text{bj}} \leq 0.01$ and $Q^2 \leq 650$ – i.e. the good fit quality shown earlier is *not* driven by the large x and Q^2 part of phase space. The choice $m_{u,d,s} = 140$ MeV for the light flavors, although popular in the literature, is not optimal and the redefinition of the Bjorken variable $x_{\text{eff}} = x_{\text{bj}}(1 + 4m_f^2/Q^2)$ actually reduces the fit quality even further. Therefore, in Figs. 18 the curves corresponding to the modified BK evolution are based on $\text{BK}_{\text{phys}}^{\text{A}}$ in which $x_{\text{eff}} = x_{\text{bj}}$. The fit to the diffractive data is fairly good.
- Asymptotic fits without the energy conservation correction ($[\text{BK}]_{\text{rc}}$ of Secs. 3 and 4 or $[\text{BK}_{140}^{\text{A}}]_{\text{rc}}$ introduced here) result in a poor fit even within the range $x_{\text{bj}} \leq 0.01$ and $Q^2 \leq 45 \text{ GeV}^2$. The fit within the wider range $Q^2 \leq 650 \text{ GeV}^2$ is a failure. The modification $m_f = 140 \rightarrow 5$ MeV improves the fit slightly, giving $\chi^2/200 = 1.65$ with $\Lambda = 52.7$ MeV and $\sigma_0 = 56.1 \text{ GeV}^{-2}$. The diffractive data are not considered.
- For the pre-asymptotic procedure $[\text{BK}_{140}^{\text{P},x_{\text{eff}}}]_{\text{rc}}$ of [45], a good fit is obtained for $x_{\text{bj}} \leq 0.01$ and $Q^2 \leq 45 \text{ GeV}^2$. The fit quality remains good when the results are extrapolated to $Q^2 = 650 \text{ GeV}^2$. The fit to the diffractive data turns out to be excellent. The ratio $\sigma_{\gamma^*p}^{\text{diff}}/\sigma_{\gamma^*p}^{\text{tot}}$ is not investigated. It is clear, however, that if both the diffractive and the inclusive cross section data are well resolved, then the same holds also for their ratio.

Close inspection of Table 7 reveals that $[\text{BK}_{140}^{\text{P},x_{\text{eff}}}]_{\text{rc}}$ shows the best χ^2 -value for diffractive data presented in this paper. The origin for this behavior can be discerned from the right panel of Fig. 18, which presents the diffractive structure functions obtained with the asymptotic and pre-asymptotic approaches. The overall shape of each component is the same in both approaches but the magnitudes are systematically smaller for pre-asymptotic case. This is especially pronounced at small β and explains the somewhat better match with the data. Unfortunately, it is the small β region where our present theoretical input, in particular the expression used to approximate the $q\bar{q}g$ contribution is deficient. Thus one cannot make any drastic conclusions based on this region: further corrections may change the results fundamentally. Our present theoretical setup is not firm enough to make use of the more detail information inherent in diffractive data to further constrain our analysis.

6 Conclusions

The main message from this study is that JIMWLK evolution at NLO allows an asymptotic fit to all HERA data below $x_{\text{bj}} \leq .02$, to both total and diffractive cross sections. This becomes possible only after *all* NLO corrections – including the energy conservation correction – are included. The energy conservation correction is the decisive ingredient in this argument: an asymptotic fit without it barely works for $Q^2 \leq 45 \text{ GeV}^2$ and fails entirely beyond. Contrary to that, with the energy conservation correction included, the fit obtained for $Q^2 \leq 45 \text{ GeV}^2$ simply *extrapolates* to a successful fit all the way up to $Q^2 = 1200 \text{ GeV}^2$, i.e. the low Q^2 range determines the result fully, over a range much larger than expected from naive BFKL based momentum space diffusion arguments. While our study does not exclude that a pre-asymptotic component is compatible with data also if one includes the energy conservation correction, such a feature is not required with present accuracy.

This observation is quite striking in that such an asymptotic fit is highly constrained and largely determined by perturbation theory: the shape of the asymptotic, pseudo-scaling correlators themselves is fully determined by the structure of the r.h.s. of the evolution equation, which is the result of a purely perturbative calculation.

The pre-asymptotic fits of [45, 46], which omit the energy conservation corrections, on the other hand, rely strongly on the fact that the evolution trajectory matched onto data has not yet reached the asymptotic domain. Both evolution speed and correlator shapes can only be matched onto data as long as many of the features of the initial conditions are not yet erased. This remains true even though the *precise* details of these initial conditions are not too strongly constrained as soon as one allows evolution trajectories that start with shapes not related to the GB-W model by evolution.

This puts an additional emphasis on the role of NLO corrections: without at least the inclusion of running coupling corrections a fit to HERA data from evolution equations is virtually impossible – it becomes *very* hard to find an initial condition that would allow us to match the full HERA range, mostly due to too fast evolution. NLO corrections bring both qualitatively new features (such as scale breaking via the running of the coupling) as well as quantitative modifications (slowdown via the energy conservation correction) that make a data fit successively easier, until, with both included, even the UV details are naturally imprinted by the evolution equation itself.

All other results are secondary to these observations, but they round out the picture and point us to what main theoretical improvements are needed to step beyond what can be done presently.

The total cross section fits work best if one assumes massless quarks – a fit with physical quark masses starts to deviate from at Q^2 values below mass of the heaviest quark included in the fit. Phenomenologically only u,d,s, and c quarks need to be considered and only the charm quark with a mass of the order of Q_s induces a strong modification. Since all these modifications are concentrated in the infrared, one should either drop this region from consideration altogether and accept the loss of fit quality in this region, or adopt non-perturbative arguments to improve the fit. One such strategy improves fit quality in the infrared by distorting how one maps the evolution trajectories onto phase space by replacing x_{bj} by some Q^2 -dependent x_{eff} . Even with rough models for x_{eff} , plausible charm fractions are achieved. *Any* refinement that aims at quantitative rather than qualitative improvements, however, must incorporate nonperturbative information from outside the scope of JIMWLK evolution.

The asymptotic fits to diffractive data use the parameters already determined via the total cross section. The formalism clearly requires contributions from NLO impact factors to fill in phase space at $\beta \rightarrow 0$. Without them a fit to data is *not* possible. While this was known empirically already from fits within the GB-W model, from an evolution perspective this becomes a consistency requirement. As soon as more than one Fock-component comes into play the details of how one models the impact parameter dependence of their scattering begins to affect fit quality: the choice of model for this non-perturbative aspect starts to affect the relative weight of these Fock-components in the associated cross section. One of the few direct experimental constraints in this respect comes from the ratio of total to diffractive cross sections, and can be accommodated easily with Gaussian profiles no matter if one uses a globally factorized form or enforces exact Casimir scaling on the other extreme.

At present, two issues hamper precision fits that involve differential cross sections with exclusive final states: the lack of complete NLO impact factors, and a consistent treatment of the impact-parameter dependence based on non-perturbative input. Only the first of these issues is sure to be

resolved in the near future, but it requires considerable effort: The full NLO corrections to JIMWLK evolution and the NLO impact factors have to be calculated. Beyond that a treatment of resummed collinear corrections needs to be formulated that eliminates any double counting issues that affect our current inclusion of the energy conservation corrections. All of these can be addressed within perturbation theory. Any progress on the impact parameter dependence requires is a different matter altogether.

Acknowledgments

J.K. and K.R. acknowledge the support from the Academy of Finland grant number 1134018. J.K. has also been supported by Jenny and Antti Wihuri Foundation and Oulu University Scholarship Foundation.

A Kinematics and common approximations

For completeness we include a brief description of the kinematical variables involved in the inclusive and diffractive deep inelastic scattering processes.

Only two independent variables are required for the process $\gamma^*(q) p(P) \rightarrow \text{anything}$, the photon virtuality Q^2 and the Bjorken variable x :

$$Q^2 := -q^2 := -(k - k')^2 \quad (\text{A.1a})$$

$$x_{\text{bj}} := \frac{Q^2}{2P \cdot q} \quad (\text{A.1b})$$

where, at leading twist, x_{bj} carries the interpretation of the momentum fraction carried by the quark inside the target that is struck by the virtual photon. The reduction to two kinematic variables holds, wherever

$$P^2/Q^2 = m_p^2/Q^2 \ll 1 \quad \text{and} \quad 1/x_{\text{bj}} \gg 1 \quad (\text{A.2})$$

to guarantee that s_{γ^*p} , the total energy squared of the $\gamma^*(q) p(P)$ -subprocess (frequently denoted W^2) can be expressed in terms of Q^2 and x_{bj} only:

$$s_{\gamma^*p} = (P + q)^2 = P^2 + 2P \cdot q - Q^2 \approx Q^2 \left(\frac{1}{x_{\text{bj}}} - 1 \right). \quad (\text{A.3})$$

The conditions (A.2) are satisfied for the majority of the experimental data.

In addition to the variables in Eqs. (A.1), two additional kinematic variables are required to describe the kinematics of the diffractive process $\gamma^*(q) p(P) \rightarrow X(M_X) p(P')$:

$$x_{\mathbb{P}} = \frac{(P - P') \cdot q}{P \cdot q} = \frac{Q^2 + M_X^2 - t}{Q^2 + s_{\gamma^*p} - m_p^2} \approx \frac{Q^2 + M_X^2}{Q^2 + s_{\gamma^*p}}, \quad (\text{A.4a})$$

$$\beta = \frac{-q^2}{2(P - P') \cdot q} = \frac{Q^2}{Q^2 + M_X^2 - t} \approx \frac{Q^2}{Q^2 + M_X^2} = \frac{x_{\text{bj}}}{x_{\mathbb{P}}}, \quad (\text{A.4b})$$

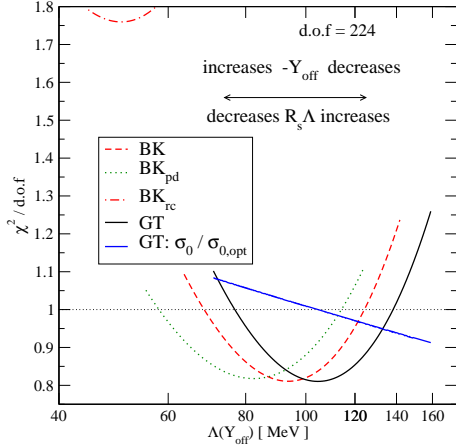


Fig. 19: χ^2/dof as a function of $\Lambda(Y_{\text{off}})$. Each evaluated χ^2/dof corresponds to a different choice of $N(Y=0, r)$ with the optimal Λ and σ_0 . The line at $\chi^2/\text{dof} = 1$ shows the extraction point of the error bars for Λ (see Figs. 7). In the asymptotic approach the offsets Y_{off} are actually negative meaning that one discards a certain number of configurations from the beginning of the evolution (typically $Y = 0$ is in the pseudo-scaling stage). The horizontal crossing line (blue) indicates the normalization dependence via $\sigma_0/\sigma_{0,\text{opt}}$ where $\sigma_{0,\text{opt}} = 55.33 \text{ GeV}^{-2}$ is the normalization of the best overall χ^2/dof . This log-linear dependence of the normalization is common to all descriptions. Note the red dot-dashed curve in the top left corner that corresponds to the data fit without the energy conservation correction.

where $x_{\mathbb{P}}$ is the fraction of the proton four momentum carried by the colorless diffractive exchange called the Pomeron. The variable β is an analogue of the x_{bj} for the diffractive system: it is the momentum fraction of the Pomeron carried by the interacting parton inside the Pomeron. The variable M_X is the invariant mass of the diffractive hadronic final state denoted by X and $t = (P - P')^2 \leq 0$ is the squared four momentum transfer. Experimentally $|t| \ll Q^2$, M_X^2 , thus t is set to zero in the above equations in addition to the proton mass m_p .

B Parameter optimization

The total error for the inclusive data is obtained by adding the systematic error quadratically to the statistical error as follows,

$$\epsilon_{\text{tot}} = \pm \sqrt{(\text{stat})^2 + \max[(\text{sys}_{\pm})^2]} \quad ; \quad \chi^2/\text{dof} = \sum_i \frac{(F_{2,\text{exper.}}^i - F_{2,\text{theor.}}^i)^2}{(\epsilon_{\text{tot}}^i)^2} / \text{dof} , \quad (\text{B.1})$$

which results in symmetric error bars. For the inclusive data this is fine since only a small number of data points have asymmetric systematic errors. However, in the case of the diffractive data the asymmetric error bars are kept due to the large differences between $|\text{sys}_+|$ and $|\text{sys}_-|$: a bias of this size in the error propagation cannot be ignored. A common χ^2 -test for the goodness of fit is applied, shown on the right in (B.1). The parameter optimizing procedure is illustrated in Fig. 19 where χ^2/dof is plotted as a function of the (correlated) parameters $\Lambda(Y_{\text{off}})$.

In some figures and tables we show the error bars for the BK based descriptions. The origin of these is the data fits and they are based on the arbitrary choice made by us. Even though there are three parameters to be fitted, it turned out that the procedure of seeking the minima of χ^2 is effectively one parameter fit: two of the parameters are correlated and the third one is just a trivial normalization. We illustrate this in Fig. 19 where χ^2/dof is plotted as a function of the correlated parameters $\Lambda(Y_{\text{off}})$. The horizontal dotted line indicates the point from where the error bars are extracted: it is *chosen to be* at $\chi^2/\text{dof} = 1$.

C Numerical implementation of the energy conservation correction

The numerical treatment of the BK- and Gaussian truncations of JIMWLK evolution is identical, an appropriate invertible map relates the two truncations for $q\bar{q}$ -dipoles [51], despite any differences encountered in the application to more general correlators. This statement in fact extends to the inclusion of the energy conservation corrections as shown in Eqs. (2.32) and (2.33). Since the actual numerical implementation used is done in BK-form, we phrase the discussion here in BK form only. With the exception of the energy conservation correction, the discretization and the numerical evaluation of the evolution equation follow [51]. Since the computational cost of a single right-hand side evaluation scales like $N_{\text{parent}} \times N_{\text{daughter}} \times N_{\text{angle}}$ (where N_i denote the number of grid points in polar coordinates), optimization of the number of discretization points is highly recommended when solving the modified BK equation²⁴.

The Y -derivative appearing on the r.h.s. of Eq. (2.32) can not be evaluated directly, to cope with it we have implemented an iterative procedure. We first write the evolution equation in finite difference form

$$\frac{S_{i+1}^k - S_i}{\delta Y} = f[S_i] - \left[\frac{f[S_{i+1}^{k-1}] - f[S_i]}{\delta Y} \right] \equiv F[S_{i+1}^{k-1}] \quad ; \quad f[S] = \text{r.h.s. of BK} , \quad (\text{C.1})$$

where the second term on the r.h.s. represents the energy conservation correction.

In (C.1), S is a function of dipole size and i labels discrete rapidity steps. k is used to label iteration steps at fixed rapidity used to deform the solution of the equation without energy conservation correction, i.e. S_{i+1}^0 , defined via

$$\frac{S_{i+1}^0 - S_i}{\delta Y} = f[S_i] \quad (\text{C.2})$$

into a solution of the full equation (C.1) at $S_{i+1} := S_{i+1}^{k \rightarrow \infty}$, where the limit $k \rightarrow \infty$ assumes convergence of the procedure. Leaving convergence issues aside for the moment, the iteration with (C.1) proceeds as follows:

$$\begin{aligned} & 1) \text{ evaluate } f[S_{i+1}^{k-1}] \quad 2) \text{ calculate } F[S_{i+1}^{k-1}] \quad 3) \text{ evolve } S_{i+1}^k = S_i^0 + \delta Y F[S_{i+1}^{k-1}] \\ & 4) \text{ return to 1) with new configuration from step 3).} \end{aligned} \quad (\text{C.3})$$

The iterative solution is accepted as the solution S_{i+1} for the next time step at some finite k , once left and right hand sides of the discretized equation (C.1) agree to some desired accuracy ϵ

$$\frac{S_{i+1}^k - S_i^0}{\delta Y} - F[S_{i+1}^k] < \epsilon , \quad (\text{C.4})$$

or equivalently $F[S_{i+1}^k] - F[S_{i+1}^{k-1}] < \epsilon$.

²⁴High computational cost of the R.H.S evaluation in the iterative procedure together with the requirement a small rapidity step δY makes the overall numerical cost of a single evolution trajectory far too large.

Note that the definition of evolution speed λ in Eq. (2.39) can be used to directly translate the accuracy criterion of Eq. (C.4) into the error implied for λ :

$$\lambda_\epsilon := 2 \int \frac{dr}{r} [\text{l.h.s}(\text{C.1}) - \text{r.h.s}(\text{C.1})] = 2 \int \frac{dr}{r} \left[\frac{S_{i+1}^k - S_i^0}{\delta Y} - F[S_{i+1}^k] \right]. \quad (\text{C.5})$$

The main issue with this procedure is that it is not stable for arbitrary shapes of the dipole function S : convergence towards (C.4) occurs only for a small number of iteration steps before the iterations start to diverge in the sense of an asymptotic series. Convergence can be improved by reducing δY (to, in general, impractical values) with the pseudo-scaling region showing the best convergence properties.

A numerical exploration reveals that it is the nonlinear region $r^2 > R_s^2(Y)$, where $N_Y(r) = 1 - S_Y(r)$ approaches one, that is least stable. The source of the instability is the nonlinearity, but the convergence properties of the asymptotic series and the obtainable accuracy can be improved by reducing the step size δY .

The manner in which δY affects convergence of the iteration procedure is slightly peculiar since the energy conservation term [in the update step (Eq. (C.3), step 3)] carries no explicit overall power of δY . However, its initial size (in the first iteration step at $k = 1$) is determined by the difference between S_i and S_{i+1}^0 as induced by (C.2). This difference is proportional to δY and thus step size imprints itself on the whole subsequent iteration procedure: convergence can be improved by reducing δY .

The main criterion for convergence therefore is the difference between S_i and S_{i+1}^0 , and this not only depends on δY itself, but also on the shape of S_i : a solution near the pseudo-scaling regime generically leads to a smaller energy conservation correction and better stability than a solution that is far from pseudo-scaling, such as the Gaussian shape used by GB-W.

Since stability is most precarious at $r^2 > R_s^2(Y)$, there is an interplay between IR regulators applied to tame the Landau pole and the convergence of the iteration procedure: Irrespective of the shape of the solutions, stability is improved whenever $\delta Y K_{\mathbf{x}\mathbf{z}\mathbf{y}} R_{\mathbf{x}\mathbf{z}\mathbf{y}}^{\text{eff}}$ is small. $R_{\mathbf{x}\mathbf{z}\mathbf{y}}^{\text{eff}}$ plays the role of an effective coupling and we find that

- the linear region $r \lesssim R_s(Y)$ is stable even with the relatively large fixed coupling $\alpha_s = 0.4$. A fast convergence occurs in all cases.
- the non-linear region $r > R_s(Y)$ is unstable even with the relatively small fixed coupling $\alpha_s = 0.2$. The case $\alpha_s = 0.4$ is already challenging and requires an impractically small step size δY . The convergence is generally slow.

For the realistic case with the full running coupling kernel (whose size is essentially determined by the size of the parent dipole r) this implies that the problems caused by the presence of the nonlinearities are exacerbated by the running of the coupling. Thus, the iteration is quick and stable at $r \lesssim R_s(Y)$ thanks to both the smallness of the kernel and absence of nonlinearities. In contrast to that, the region $r > R_s(Y)$ is difficult (especially at low rapidities) since the kernel is large and the equation is dominated by non-linear effects.

In practical terms, the convergence properties of the iteration procedure preclude the use of (C.3) away from the asymptotic line. We have applied (C.3) in the pseudo-scaling region, iteratively

reducing δY to push the step of minimal error k_{\min} to larger k and minimize the error in a brute force approach to set a baseline. As is typical with iteration procedures, convergence properties can be strongly affected by a modification of the iteration procedure. As a compromise between speed and accuracy near the pseudo-scaling region we have amended steps 1) – 2) of (C.3) by a set of re-weighting steps:

$$\begin{aligned}
& 3) \ S_{i+1}^k = S_i^0 + \delta Y F[S_{i+1}^{k-1}] \quad ; \quad G_{i+1}^k = 2S_{i+1}^{k-1} - S_{i+1}^k \\
& 4) \text{ evaluate } f[S_{i+1}^k] \text{ and calculate } F[S_{i+1}^k] \\
& 5) \ S_{i+1}^{k+1} = S_i^0 + \delta Y F[S_{i+1}^k] \quad ; \quad G_{i+1}^{k+1} = 2S_{i+1}^k - S_{i+1}^{k+1} \\
& 6) \text{ calculate } \mathcal{W}_{i+1}^k = S_{i+1}^k - \Delta_s (G_{i+1}^k - S_{i+1}^k) / (\Delta_g - \Delta_s) \\
& \text{where } \Delta_s = \sum_r (S_{i+1}^{k+1} - S_{i+1}^k) \quad ; \quad \Delta_g = \sum_r (G_{i+1}^{k+1} - G_{i+1}^k) \\
& 7) \text{ return to 1) with weighted } \mathcal{W}_{i+1}^k \text{ of 6) taking the place of } S_{i+1}^k
\end{aligned} \tag{C.6}$$

The procedure terminates when $|\Delta_s/\Delta_g| - 1| < \epsilon$. Then further iterations of \mathcal{W}_{i+1}^k do not lead to any improvement²⁵ and one proceeds with S_{i+1}^k in step 4). Initially (typically) $|\Delta_g| \gg |\Delta_s|$ and so G_{i+1}^k gets weighted less than S_{i+1}^k in \mathcal{W}_{i+1}^k .

The upshot is high precision at small r after few iterations already at $\delta Y = 0.0025$, at the price of comparatively limited precision in the large r part and with very few iteration steps. The procedure always terminates but can, without additional modifications, not exceed the precision indicated in Figs. 20 and 21.

We observe that the unmodified method (C.3) at $k = 1$, i.e. the crudest approximation, greatly underestimates λ whereas the first iteration $k = 2$ overestimates it, as seen in the left panel of Fig. 20.

From $k = 3$, the brute force solution to (C.3) is more accurate at large r than the approximative solution (C.6), however, this solution is increasingly difficult to obtain: in order to keep the iteration stable one has to reduce the change in shape of S with Y by working near the pseudo-scaling region,²⁶ and *in addition* needs to employ a very small rapidity step $\delta Y \lesssim 10^{-3}$. To stabilize (C.3) also for $k = 4, 5$ requires $\delta Y \lesssim 10^{-4}, 10^{-5}$ respectively which increases the numerical effort prohibitively. This is illustrated in the right panel of Fig. 20 which compares the three step iteration of (C.3) with the result of the re-weighted result from (C.6) obtained with almost an order of magnitude smaller CPU time. Fig. 21 shows the size of the energy conservation correction in the different cases. For both Figs. 20 and 21, the configurations are selected from the Y - or $R_s\Lambda$ -range typically used in the data fits²⁷.

The induced error of evolution speed according to (C.5) is shown in Table 8 for three samples corresponding to $R_s\Lambda = \{0.3, 0.2, 0.1\}$. The region $r \leq R_s$ is fairly accurate in all cases and especially the case $k = 3$ is close to the accurate solution even at $r > R_s$. The peak values (Figs. 20, middle) correspond to $S_{Y,xy} \approx 0.2$. As seen, the iterations $k = 4, 5$ do not bring any

²⁵The condition $\Delta_s \approx -\Delta_g$ yields $\mathcal{W}_{i+1}^k \approx (S_{i+1}^k + G_{i+1}^k)/2$ regardless of $|\Delta_{s,g}| \ll 1$ which obviously is a bad solution for any r .

²⁶This requires some numerical optimization of its own.

²⁷In the data fits $\Lambda \approx 0.1$ GeV and so $R_s\Lambda = \{0.1, 0.2, 0.3\} \rightarrow R_s \approx \{1, 2, 3\}$ GeV⁻¹.

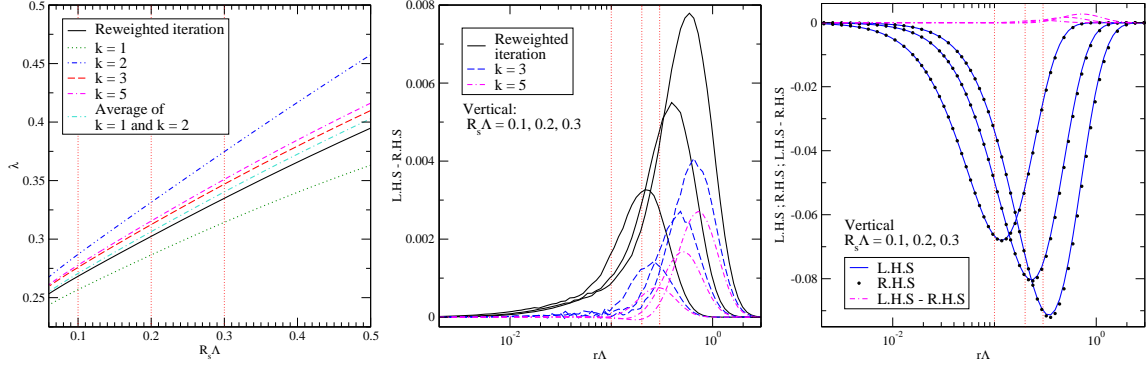


Fig. 20: **Left:** λ : a comparison of iterations $k = 1, 2, 3, 5$ of (C.3), the average of $k = 1, 2$ and the procedure (C.6). **Middle:** Accuracy of the solution from (C.4) as a function of r for $R_s\Lambda = \{0.1, 0.2, 0.3\}$ (corresponds $R_s \approx \{1, 2, 3\} \text{ GeV}^{-1}$). **Right:** The solution of the case $k = 5$ is, for all practical purposes, accurate but the requirement of computing time is huge. L.H.S. and R.H.S. refer to the discretized evolution equation (C.1)

	δY	λ_ϵ	$ \lambda_\epsilon(r \leq R_s) /\lambda_\epsilon$
reweighted	0.0025	$\{0.025, 0.018, 0.011\}$	$\{0.256, 0.250, 0.245\}$
brute force $k = 3$	10^{-3}	$\{0.010, 0.007, 0.003\}$	$\{0.106, 0.092, 0.079\}$
brute force $k = 5$	10^{-5}	$\{5.824, 3.541, 1.561\} \times 10^{-3}$	$\{6.437, 0.448, 4.590\} \times 10^{-3}$

Table 8: The error estimates for λ based on (C.5) for different iteration procedures (first column). The numerical cost is measured by δY . Most of the error comes from $|\mathbf{r}| > R_s(Y)$ (last column). The values correspond to correlation lengths $R_s\Lambda = \{0.3, 0.2, 0.1\}$.

essential improvement and for $k > 5$ the iteration procedure turned out to be highly unstable²⁸ at $r > R_s$ for any relevant δY .

It should be noted that a convergent solution in the offset region²⁹ is easier to obtain with some other, say stronger, regulator but, however, inside the actual fit range this kind of modification does not bring any improvement.

D Tools to efficiently address diffractive cross sections

A lot of simplifications used in earlier treatments with simple height one box profiles can be carried over to more general profile models if the \mathbf{r} and \mathbf{r}' dependence in the \mathbf{b} -integrated product of dipole amplitudes factorizes according to³⁰

$$\int d^2\mathbf{b} N_{\mathcal{R},Y}^*(\mathbf{r}, \mathbf{b}) N_{\mathcal{R},Y}(\mathbf{r}', \mathbf{b}) = \sigma_{\mathcal{R}}^{(2)} [N_{\mathcal{R},Y}^{(2)}(\mathbf{r})]^* N_{\mathcal{R},Y}^{(2)}(\mathbf{r}') , \quad (\text{D.1})$$

or if such factorized behavior is a good approximation to the full result c.f. (4.10).

²⁸The iteration seems to be approaching (alternatingly) the fixed point but the requirement of a very small δY makes the brute force method impractical.

²⁹The region of large running coupling that is cut off by the parameter Y_{off} .

³⁰In (D.1) the constants $\sigma_{\mathcal{R}}^{(2)}$ are adjusted such that $N_{\mathcal{R},Y}^{(2)}(\mathbf{r}) \xrightarrow{r \rightarrow \infty} 1$ for convenience.

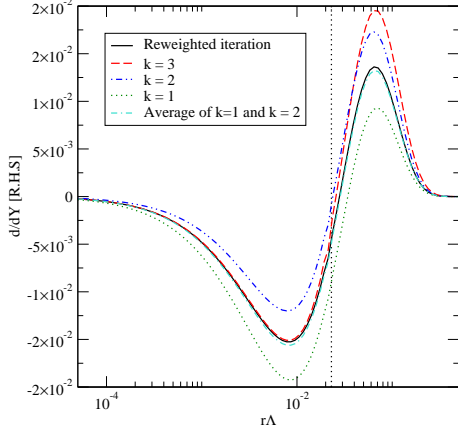


Fig. 21: The energy conservation correction $(f[S_{i+1}^{k-1}] - f[S_i^0])/\delta Y$ for the cases $k = 1, 2, 3$ of (C.3), an average of $k = 1, 2$ and the procedure (C.6). The latter approximately coincides with the average of $k = 1, 2$ for all r and thus the easiest way to get a fairly good solution is taking the average of $k = 1, 2$. A sharp slowly converging peak at large r is a consequence of the non-linearity since it appears even with a relatively small fixed coupling $\alpha_s = 0.2$. The running coupling, being large at these scales, gives rise to the additional stability problems.

Wherever (D.1) holds, a momentum space variant of the cross sections of (4.1) offers an efficient route to perform the data fits. It allows to pre-calculate most of the numerical integrals in the final expressions for the diffractive structure function and speeds up the parameter seeking process tremendously. The tools to do so are collected in the remainder of this section.

D.1 Diffractive cross sections in momentum space

The momentum space expressions for the three terms in Eq. (4.1) read

$$\begin{aligned}
\frac{d\sigma_{q\bar{q},T}^{\gamma^*A \rightarrow Xp}}{d\beta} &= \frac{(2\pi)^2 \alpha_{\text{em}}}{Q^2 \beta} \sum_f e_f^2 \int_{z_f^-}^{z_f^+} dz \frac{Q^2(z^2 + \bar{z}^2) N_c}{32\pi \left(\bar{\beta} - \frac{m_f^2}{\bar{\beta}} \right)} \\
&\times \int d^2 \mathbf{b} \left[\int \frac{dq^2}{q^2} \mathcal{N}_Y^{q\bar{q}}(q, b) \left(1 - 2\beta - \frac{2m_f^2}{\bar{\beta}} - \frac{(1-2\beta) \frac{z\bar{z}Q^2}{\bar{\beta}} - (q^2 + 2m_f^2)}{\sqrt{(\frac{z\bar{z}Q^2}{\bar{\beta}} + q^2)^2 - 4(\bar{\beta} \frac{z\bar{z}Q^2}{\bar{\beta}} - m_f^2) q^2}} \right) \right]^2 \\
&+ \frac{(2\pi)^2 \alpha_{\text{em}}}{Q^2 \beta} \sum_f e_f^2 \int_{z_f^-}^{z_f^+} dz \frac{\beta m_f^2 N_c}{8\pi z \bar{z}} \\
&\times \int d^2 \mathbf{b} \left[\int \frac{dq^2}{q^2} \mathcal{N}_Y^{q\bar{q}}(q, b) \left(1 - \frac{\frac{z\bar{z}Q^2}{\bar{\beta}}}{\sqrt{(\frac{z\bar{z}Q^2}{\bar{\beta}} + q^2)^2 - 4(\bar{\beta} \frac{z\bar{z}Q^2}{\bar{\beta}} - m_f^2) q^2}} \right) \right]^2 \tag{D.2a}
\end{aligned}$$

$$\begin{aligned}
\frac{d\sigma_{q\bar{q},L}^{\gamma^*A \rightarrow Xp}}{d\beta} &= \frac{(2\pi)^2 \alpha_{\text{em}}}{Q^2 \beta} \sum_f e_f^2 \int_{z_f^-}^{z_f^+} dz \frac{\beta 4Q^2(z\bar{z})^2 N_c}{8\pi z \bar{z}} \\
&\times \int d^2 \mathbf{b} \left[\int \frac{dq^2}{q^2} \mathcal{N}_Y^{q\bar{q}}(q, b) \left(1 - \frac{\frac{z\bar{z}Q^2}{\bar{\beta}}}{\sqrt{(\frac{z\bar{z}Q^2}{\bar{\beta}} + q^2)^2 - 4(\bar{\beta} \frac{z\bar{z}Q^2}{\bar{\beta}} - m_f^2) q^2}} \right) \right]^2 \tag{D.2b}
\end{aligned}$$

$$\begin{aligned}
\frac{d\sigma_{q\bar{q}g,T}^{\gamma^*A \rightarrow Xp}}{d\beta} \Big|_{\text{LL}Q^2} &= \frac{(2\pi)^2 \alpha_{\text{em}}}{Q^2 \beta} \frac{\alpha_s \beta C_f N_c}{32\pi^2} \sum_f e_f^2 \int_{\beta}^1 \frac{dz}{z^2 \bar{z}^2} \left[\left(1 - \frac{\beta}{z} \right)^2 + \left(\frac{\beta}{z} \right)^2 \right] \int_0^{Q^2} dk^2 \ln \left(\frac{Q^2}{k^2} \right) \\
&\times \int d^2 \mathbf{b} \left[\int \frac{dq^2}{q^2} \mathcal{N}_Y^{gg}(q, b) \left(z^2 + \bar{z}^2 + \frac{q^2}{k^2} - \frac{(q^2 - k^2(1-2z))^2 + 2z\bar{z}k^4}{k^2 \sqrt{(q^2 + k^2)^2 - 4(1-z)q^2 k^2}} \right) \right]^2 \tag{D.2c}
\end{aligned}$$

Where we have used the shorthand expressions $\bar{z} = 1 - z$ and $z_f^\pm = \frac{1}{2}(1 \pm \sqrt{1 - 4m_f^2/M_X^2})$.

In the massless limit, these expressions match up with their counterparts in [31, 69]³¹ once the corresponding b -profiles³² have been inserted and the impact parameter integral has been carried out.

The cumbersome expressions in brackets arise from integrations over the orientation of \mathbf{q} of rather simple Fourier expressions for the McDonald K_i appearing in the coordinate space variants. The solutions of the evolution equations enter via $\mathcal{N}_Y^{\mathcal{R}}(q, b)$, which are determined from the coordinate

³¹Use $\alpha_s \mathcal{F}(x, \mathbf{q}) = \frac{N_c \sigma_0}{4\pi} \mathcal{N}_{\bar{q}q, \ln(1/x)}(\mathbf{q})$ to translate the $q\bar{q}$ expressions into those of [31, 69]. The $q\bar{q}g$ -term in the original literature suffers from an additional incorrect rescaling by a factor of $(C_f/N_c)^2$ and otherwise substitute $\mathcal{N}_{q\bar{q}}$ for \mathcal{N}_{gg} [53, 71].

³²Where ever these papers make use of the $d\sigma/dt|_{t=0}$ in their expressions one *must* use factorized Gaussian profiles for this procedure, the corresponding expressions are valid only in this case.

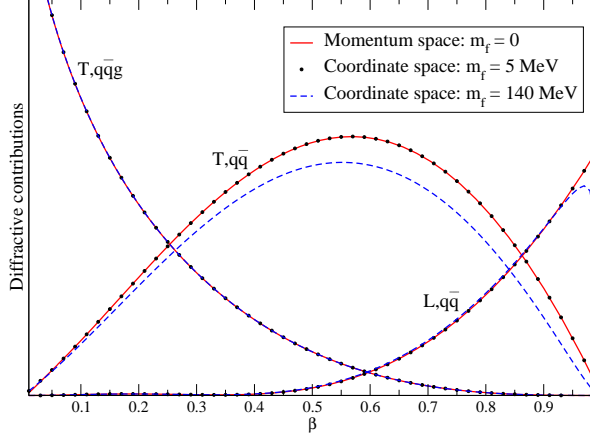


Fig. 22: A comparison between the momentum and coordinate space formulae of the diffractive final states. Here, the same dipole input (a quark dipole) is used in all contributions and the parameters B_d and α_s are set to one. In the case of the coordinate space equations, three light flavors with equal masses $m_{u,d,s} = 5, 140$ MeV are considered whereas for the momentum space equations $m_f = 0$ (coincides perfectly with the case $m_{u,d,s} = 5$ MeV). The curves are calculated with fixed $Q^2 = 15 \text{ GeV}^2$ and $x_{\mathbb{P}} = 0.001$.

space dipole amplitudes $N_Y^{\mathcal{R}}(r, b)$. These are *not* related by a direct Fourier transform, see Sec. D.2 for definitions and properties.

The benefit of using the momentum space forms of the equations shows up in the numerical implementation: the required rapidity range of \mathcal{N} can be pre-calculated whereas the coordinate space expressions have highly oscillatory integrands in which no part of the nested integrals can be pre-calculated.

We have cross-checked our results using both coordinate and momentum space variants numerically. A direct comparison of the quark contributions at different masses gives an idea of how they affect the cross sections: non-zero quark masses reduce $x_{\mathbb{P}} F_{T,q\bar{q}}^{D(3)}$ whereas $x_{\mathbb{P}} F_{L,q\bar{q}}^{D(3)}$ is practically unaffected. In any case, it is expected that considering the non-zero quark masses would only lead to a small rescaling of the normalizations in Eq. (4.1)

D.2 Integral transformations for dipole amplitudes

The key ingredient is the non-standard “Fourier”-transform of the dipole amplitude (here adapted to the azimuthally symmetric forward case relevant in conjunction with (4.10))

$$N_Y^{\mathcal{R}}(\mathbf{r}) = \int \frac{d^2 \mathbf{q}}{q^2} (1 - e^{i\mathbf{r} \cdot \mathbf{q}}) \mathcal{N}_Y^{\mathcal{R}}(\mathbf{q}) \quad (\text{D.3})$$

($\mathcal{R} = \bar{q}q, gg$ etc. labels the representation; see [62] for a broader exposition on the structure of the exponentials) together with its inverse³³

$$\mathcal{N}_Y^{\mathcal{R}}(\mathbf{q}) = \left(\frac{d}{d \ln(q^2)} \right)^2 \phi_{\mathcal{R},Y}(\mathbf{q}) \quad \text{where} \quad \phi_{\mathcal{R},Y}(\mathbf{q}) := \frac{1}{\pi^2} \int \frac{d^2 \mathbf{r}}{r^2} e^{-i\mathbf{r} \cdot \mathbf{q}} N_Y^{\mathcal{R}}(\mathbf{r}) . \quad (\text{D.4})$$

Note that \mathcal{N} as used here is dimensionless just as the dipole amplitudes. This differs from the analogous quantities in [62] or in the work of GB-W. We note that \mathcal{N}/q^2 is normalized to the

³³This can be derived via an inverse Mellin transform [74, 75] (see [76] for a step by step exposition).

saturation value of the dipole amplitude:

$$\int \frac{d^2 \mathbf{q}}{q^2} \mathcal{N}_Y^{\mathcal{R}}(\mathbf{q}) = N_Y^{\mathcal{R}}(|\mathbf{r}| \rightarrow \infty) ; \quad (\text{D.5})$$

correspondingly, $S_Y^{\mathcal{R}}(\mathbf{r}) = \int \frac{d^2 \mathbf{q}}{q^2} e^{i\mathbf{r} \cdot \mathbf{q}} \mathcal{N}_Y^{\mathcal{R}}(\mathbf{q})$. Eq. (D.5) provides a stringent check for our numerical tools, which also faithfully resolve the chain of transformations $N_Y^{\mathcal{R}}(\mathbf{r}) \rightarrow \phi_Y^{\mathcal{R}}(\mathbf{q}) \rightarrow \mathcal{N}_Y^{\mathcal{R}}(\mathbf{q})$ for the GB-W model, where all the steps can be determined analytically from Eq. (D.4).

The chain of transformations $N_Y^{\mathcal{R}}(\mathbf{r}) \rightarrow \phi_Y^{\mathcal{R}}(\mathbf{q}) \rightarrow \mathcal{N}_Y^{\mathcal{R}}(\mathbf{q})$ is illustrated in Fig. 23, quark and gluon dipoles are compared in Fig. 24.

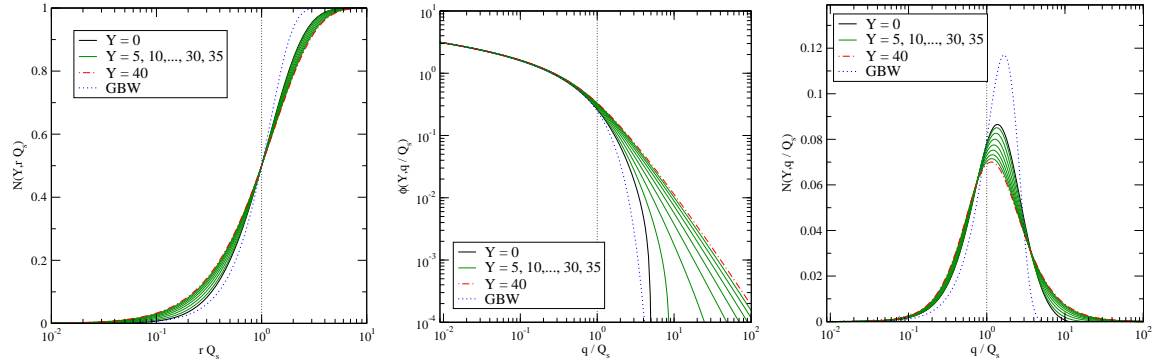


Fig. 23: From left to right: $N_Y^{q\bar{q}}(\mathbf{r}) \rightarrow \phi_Y^{q\bar{q}}(\mathbf{q}) \rightarrow \mathcal{N}_Y^{q\bar{q}}(\mathbf{q})$ transformations as a function of scaling variables rQ_s and l_t/Q_s . The vertical lines indicate the position of the saturation scale $Q_s = 1/R_s$.

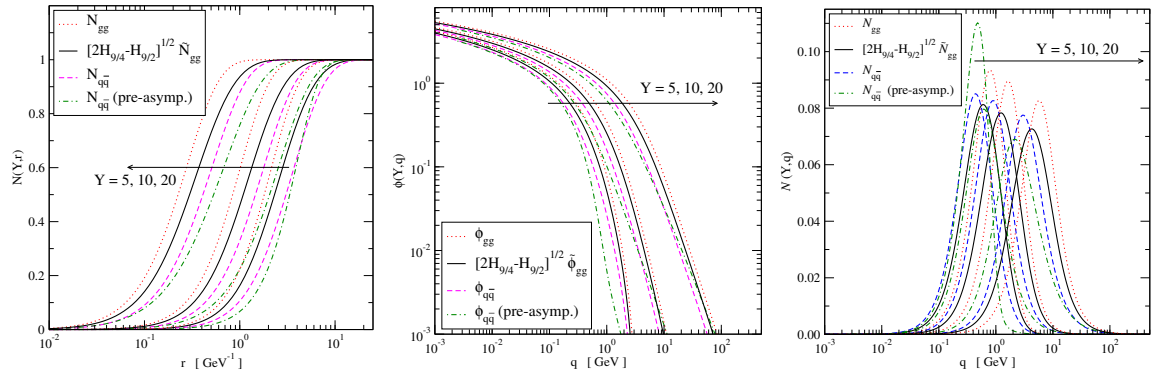


Fig. 24: A comparison of the quark and gluon dipole inputs. From left to right: $N_Y^{\mathcal{R}}(\mathbf{r}) \rightarrow \phi_Y^{\mathcal{R}}(\mathbf{q}) \rightarrow \mathcal{N}_Y^{\mathcal{R}}(\mathbf{q})$. Green dot-dashed curves are for the pre-asymptotic fit scenario studied in Sec. 5 (starting from $Y = \ln(1/0.01) \approx 4.61$).

E Consistency checks

With the importance of NLO contributions firmly established, one should at least attempt to understand in which sense the results obtained from (2.32) or (2.33) are stable against modifications. One obvious modification is a simple numerical check for stability against higher order corrections such as higher order running coupling contributions. While this has already been discussed in the quite sophisticated framework of renormalon corrections in [54] a brief numerical check on the quantitative impact of such corrections on a data fit gives an alternative ballpark impression of their impact.

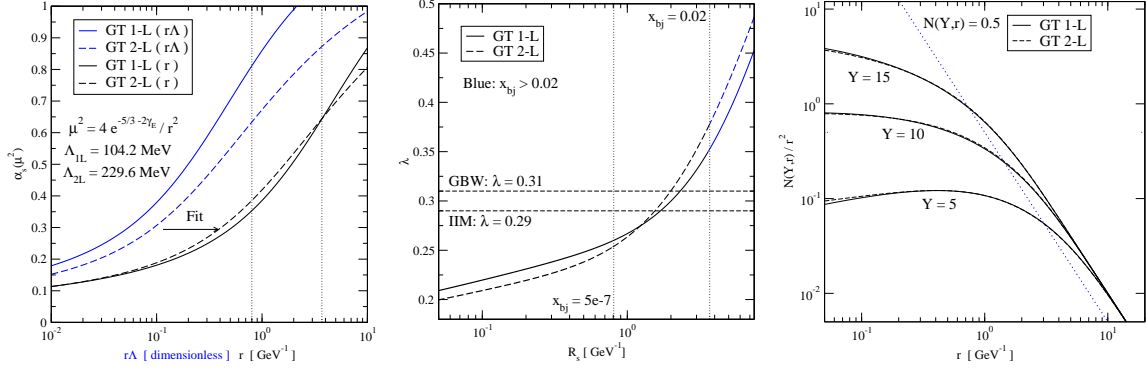


Fig. 25: The slowdown due to the running coupling is stable against higher order corrections. The vertical lines bracket the R_s -range of HERA. **Left:** One loop and two loop running couplings coincide after readjusting Λ . **Middle:** Evolution speeds differ only slightly in the HERA x -range. **Right:** The dipole correlators (magnified by r^{-2}) agree after adjusting the fit parameters with only tiny differences remaining at short distances.

The fit is repeated by using the two loop running coupling

$$\alpha_s^{2L}(\mu) = \frac{4\pi}{\beta_0} \left[\frac{1}{\ln(\mu^2/\Lambda^2)} - \frac{\beta_1}{\beta_0^2} \frac{\ln(\ln(\mu^2/\Lambda^2))}{\ln^2(\mu^2/\Lambda^2)} \right] ; \quad \beta_0 = 9, \quad \beta_1 = 64 \quad \text{for } N_f = 3, \quad (\text{E.1})$$

which was also regulated by the APT method³⁴. Once the optimal $\Lambda_{1L,2L}$ and the corresponding rapidity offsets are known, the one and two loop couplings approximately coincide (see Fig.25, left). Furthermore, the evolution speeds (Fig. 25, middle) and the dipole correlators (Fig. 25, right) coincide around $Y = 10$, which is a direct consequence of the fact that most data points are located around this rapidity. The fit quality is not essentially affected, see Table 9.

The sensitivity to the infrared regulator is investigated. Whereas the APT regulator ((E.2), right) offers a smooth crossing over the Landau pole with the limit $\alpha_s^{APT}(0) = 4\pi/\beta_0$, for instance, adding of a constant inside the logarithm as follows,

$$\alpha_s^{alt}(\mu) = \frac{4\pi}{\beta_0} \frac{1}{\ln(\mu^2/\Lambda^2 + \#)} \quad ; \quad \alpha_s^{APT}(\mu) = \frac{4\pi}{\beta_0} \left(\frac{1}{\ln(\mu^2/\Lambda^2)} - \frac{1}{\mu^2/\Lambda^2 - 1} \right), \quad (\text{E.2})$$

results in a steeply increasing (for not too large $\#$) coupling near the Landau pole. As seen in Fig. 26, the effect induced by this type of modification extends to the actual fit range. Despite the

³⁴In practice, $\alpha_s^{2L}(\mu)$ must be evaluated by using a spectral integral representation, see [63, 64].

GT ; $x_{bj} \leq 0.02$	$Q^2 \leq$	dof	χ^2/dof	$\Lambda_{1L,2L}$ [MeV]	σ_0 [GeV $^{-2}$]	$\Lambda_{2L}/\Lambda_{1L}$
One loop	45 GeV 2	224	0.81	104.7	55.33	
One loop	1200 GeV 2	295	0.80	104.2	55.24	
Two loop	45 GeV 2	224	0.81	213.8	55.41	2.04
Two loop	1200 GeV 2	295	0.83	229.6	54.59	2.20

Table 9: Fit results for GT with two loop running coupling. The quark masses are $m_{u,d,s} = 5$ MeV.

BK ; $x_{bj} \leq 0.02$	$Q^2 \leq$	dof	χ^2/dof	Λ [MeV]	σ_0 [GeV $^{-2}$]
regulator:	45 GeV 2	224	0.81	93.7	55.05
APT	1200 GeV 2	295	0.80	97.7	54.50
regulator:	45 GeV 2	224	0.96	111.7	56.57
alt.	1200 GeV 2	295	1.01	102.9	57.24

Table 10: Fit results corresponding the regulators in Eq. (E.2). The case of the alternative regulator corresponds to the choice $\# = \exp(1)$ with the limit $\alpha_s^{alt}(0) = 4\pi/\beta_0$.

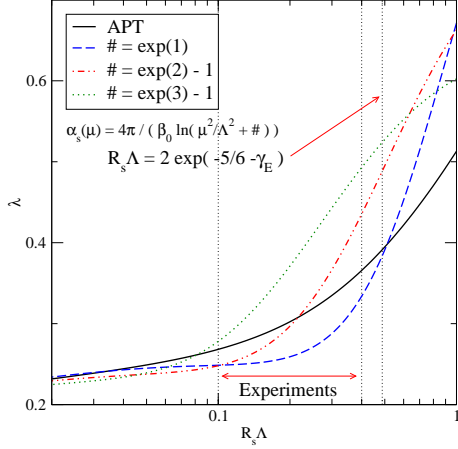


Fig. 26: A comparison of the alternative regulators and APT regulator presented in Eq. (E.2). The effect on the evolution speed is not what one would naively expect it to be: on the one hand, a weaker regulator with a larger effective running coupling R_{eff} produces a faster evolution speed at the pre-asymptotic stage than the APT (as it should) but suddenly drops below it inside the fit range $R_s \Lambda \approx 0.1 - 0.4$. On the other hand, a stronger regulator, and hence a smaller R_{eff} , extends the pre-asymptotic stage remarkably which eventually leads to a faster evolution speed inside the fit range. The evolution with the APT regulator matches the best with the data. The length scale $R_s \Lambda = 2e^{-5/6 - \gamma_E}$ indicates the location of the Landau pole.

deviating evolution speeds, all cases shown can resolve the data with a good $\chi^2/\text{dof} \lesssim 1$. However, as seen in Table 10, the evolution with the APT regulator yields the best fit to the data.

The last check concerns the solution of the modified BK equation introduced in App. C. The fit with $m_{u,d,s} = 5$ MeV is repeated by using more accurate solutions for the modified BK equation, i.e. the cases $k = 3, 5$ of the procedure (C.3). The results of these fits are presented in Table 11. In both cases χ^2/dof is increased if compared with the reweighted case shown in the first row but,

BK ; $x_{bj} \leq 0.02$	$\chi^2/295$	Λ [MeV]	σ_0 [GeV $^{-2}$]
reweighted, Eq. (C.6)	0.80	97.7	55.50
$k = 3$, Eq. (C.3)	0.86	98.2	53.89
$k = 5$, Eq. (C.3)	0.95	95.0	54.18

Table 11: Fit results corresponding to the different iterative solutions presented in App. C. In all cases, a wider data range $Q^2 \leq 1200$ GeV 2 is considered.

however, remains below one. The parameters Λ and σ_0 are altered as well but are still in the same ballpark as the ones shown in the first row. The experimental data seem to favor slower evolution speed obtained by the approximate solution.

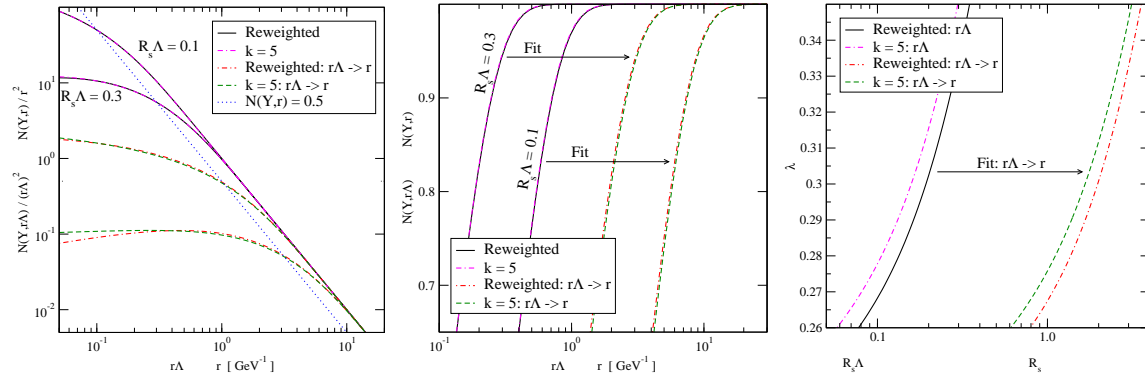


Fig. 27: Left and Middle: A comparison of $N_Y^{q\bar{q}}(r)$ of the applied and accurate solution. For a fixed $R_s\Lambda$, there is no visible difference between the shapes of the correlators. The same feature is approximately preserved after the physical scales are determined by the data fits. The correlators are extracted from $R_s\Lambda = \{0.1, 0.3\}$. The crossing blue dotted line is indicating the saturation condition $N_Y^{q\bar{q}}(r) = 0.5$. **Right:** The deviation between the corresponding evolution speeds remains after readjusting Λ . Thus, the reason for the deviation in χ^2/dof is a slightly different energy dependencies of the saturation scales rather than the actual shapes of the correlators.

To summarize, the fits based on the modified BK/GT evolution are stable against a large variety of modifications. As seen in Fig. 26, the biggest uncertainty to *evolution* clearly emerges from the infrared regulator.

References

- [1] L. V. Gribov, E. M. Levin, and M. G. Ryskin, *Semihard processes in QCD*, *Phys. Rept.* **100** (1983) 1–150. [1](#)
- [2] A. H. Mueller and J.-w. Qiu, *Gluon recombination and shadowing at small values of x* , *Nucl. Phys.* **B268** (1986) 427. [1](#)
- [3] A. H. Mueller, *Soft gluons in the infinite momentum wave function and the BFKL pomeron*, *Nucl. Phys.* **B415** (1994) 373–385. [1](#), [2.3.1](#)
- [4] A. H. Mueller and B. Patel, *Single and double BFKL pomeron exchange and a dipole picture of high-energy hard processes*, *Nucl. Phys.* **B425** (1994) 471–488, [[hep-ph/9403256](#)]. [1](#)
- [5] A. H. Mueller, *Unitarity and the BFKL pomeron*, *Nucl. Phys.* **B437** (1995) 107–126, [[hep-ph/9408245](#)]. [1](#)
- [6] L. D. McLerran and R. Venugopalan, *Gluon distribution functions for very large nuclei at small transverse momentum*, *Phys. Rev.* **D49** (1994) 3352–3355, [[hep-ph/9311205](#)]. [1](#)

- [7] L. D. McLerran and R. Venugopalan, *Computing quark and gluon distribution functions for very large nuclei*, *Phys. Rev.* **D49** (1994) 2233–2241, [[hep-ph/9309289](#)]. 1
- [8] L. D. McLerran and R. Venugopalan, *Gluon distribution functions for very large nuclei at small transverse momentum*, *Phys. Rev.* **D49** (1994) 3352–3355, [[hep-ph/9311205](#)]. 1
- [9] L. D. McLerran and R. Venugopalan, *Green’s functions in the color field of a large nucleus*, *Phys. Rev.* **D50** (1994) 2225–2233, [[hep-ph/9402335](#)]. 1
- [10] Y. V. Kovchegov, *Non-abelian Weizsaecker-Williams field and a two- dimensional effective color charge density for a very large nucleus*, *Phys. Rev.* **D54** (1996) 5463–5469, [[hep-ph/9605446](#)]. 1
- [11] Y. V. Kovchegov, *Quantum structure of the non-abelian Weizsaecker-Williams field for a very large nucleus*, *Phys. Rev.* **D55** (1997) 5445–5455, [[hep-ph/9701229](#)]. 1
- [12] J. Jalilian-Marian, A. Kovner, L. D. McLerran, and H. Weigert, *The intrinsic glue distribution at very small x* , *Phys. Rev.* **D55** (1997) 5414–5428, [[hep-ph/9606337](#)]. 1
- [13] J. Jalilian-Marian, A. Kovner, A. Leonidov, and H. Weigert, *The BFKL equation from the Wilson renormalization group*, *Nucl. Phys.* **B504** (1997) 415–431, [[hep-ph/9701284](#)]. 1
- [14] J. Jalilian-Marian, A. Kovner, A. Leonidov, and H. Weigert, *The Wilson renormalization group for low x physics: Towards the high density regime*, *Phys. Rev.* **D59** (1999) 014014, [[hep-ph/9706377](#)]. 1
- [15] J. Jalilian-Marian, A. Kovner, and H. Weigert, *The Wilson renormalization group for low x physics: Gluon evolution at finite parton density*, *Phys. Rev.* **D59** (1999) 014015, [[hep-ph/9709432](#)]. 1
- [16] J. Jalilian-Marian, A. Kovner, A. Leonidov, and H. Weigert, *Unitarization of gluon distribution in the doubly logarithmic regime at high density*, *Phys. Rev.* **D59** (1999) 034007, [[hep-ph/9807462](#)]. 1
- [17] A. Kovner, J. G. Milhano, and H. Weigert, *Relating different approaches to nonlinear QCD evolution at finite gluon density*, *Phys. Rev.* **D62** (2000) 114005, [[hep-ph/0004014](#)]. 1
- [18] H. Weigert, *Unitarity at small Bjorken x* , *Nucl. Phys.* **A703** (2002) 823–860, [[hep-ph/0004044](#)]. 1, 2.3.1
- [19] E. Iancu, A. Leonidov, and L. D. McLerran, *Nonlinear gluon evolution in the color glass condensate. I*, *Nucl. Phys.* **A692** (2001) 583–645, [[hep-ph/0011241](#)]. 1
- [20] E. Ferreiro, E. Iancu, A. Leonidov, and L. McLerran, *Nonlinear gluon evolution in the color glass condensate. II*, *Nucl. Phys.* **A703** (2002) 489–538, [[hep-ph/0109115](#)]. 1
- [21] Y. V. Kovchegov, *Small- x F_2 structure function of a nucleus including multiple pomeron exchanges*, *Phys. Rev.* **D60** (1999) 034008, [[hep-ph/9901281](#)]. 1, 2.3.1
- [22] Y. V. Kovchegov, *Unitarization of the BFKL pomeron on a nucleus*, *Phys. Rev.* **D61** (2000) 074018, [[hep-ph/9905214](#)]. 1

- [23] I. Balitsky, *Operator expansion for high-energy scattering*, *Nucl. Phys.* **B463** (1996) 99–160, [[hep-ph/9509348](#)]. 1, 2.3.1
- [24] I. Balitsky, *Operator expansion for diffractive high-energy scattering*, [hep-ph/9706411](#). 1, 2.3.1
- [25] I. Balitsky, *Factorization and high-energy effective action*, *Phys. Rev.* **D60** (1999) 014020, [[hep-ph/9812311](#)]. 1
- [26] E. Iancu and R. Venugopalan, *The color glass condensate and high energy scattering in QCD*, [hep-ph/0303204](#). 1
- [27] H. Weigert, *Evolution at small x_{bj} : The Color Glass Condensate*, *Prog. Part. Nucl. Phys.* **55** (2005) 461–565, [[hep-ph/0501087](#)]. 1, 2.3.1
- [28] J. Jalilian-Marian and Y. V. Kovchegov, *Saturation physics and deuteron gold collisions at rhic*, *Prog. Part. Nucl. Phys.* **56** (2006) 104–231, [[hep-ph/0505052](#)]. 1
- [29] F. Gelis, E. Iancu, J. Jalilian-Marian, and R. Venugopalan, *The Color Glass Condensate*, [1002.0333](#). 1
- [30] K. Golec-Biernat and M. Wüsthoff, *Saturation effects in deep inelastic scattering at low Q^2 and its implications on diffraction*, *Phys. Rev.* **D59** (1999) 014017, [[hep-ph/9807513](#)]. 1, 3.3
- [31] K. Golec-Biernat and M. Wüsthoff, *Saturation in diffractive deep inelastic scattering*, *Phys. Rev.* **D60** (1999) 114023, [[hep-ph/9903358](#)]. 1, 4.1, D.1, 31
- [32] A. M. Stasto, K. Golec-Biernat, and J. Kwiecinski, *Geometric scaling for the total γ^*p cross-section in the low x region*, *Phys. Rev. Lett.* **86** (2001) 596–599, [[hep-ph/0007192](#)]. 1
- [33] F. Caola and S. Forte, *Geometric Scaling from GLAP evolution*, *Phys. Rev. Lett.* **101** (2008) 022001, [[0802.1878](#)]. 1
- [34] **ZEUS** Collaboration, J. Breitweg *et. al.*, *Measurement of the proton structure function F_2 at very low Q^2 at HERA*, *Phys. Lett.* **B487** (2000) 53–73, [[hep-ex/0005018](#)]. 1, 1, 9, 10
- [35] **ZEUS** Collaboration, J. Breitweg *et. al.*, *ZEUS results on the measurement and phenomenology of F_2 at low x and low Q^2* , *Eur. Phys. J.* **C7** (1999) 609–630, [[hep-ex/9809005](#)]. 1, 1, 9, 10
- [36] **ZEUS** Collaboration, S. Chekanov *et. al.*, *Measurement of the neutral current cross section and F_2 structure function for deep inelastic $e+p$ scattering at HERA*, *Eur. Phys. J.* **C21** (2001) 443–471, [[hep-ex/0105090](#)]. 1, 1, 9, 10
- [37] **ZEUS** Collaboration, M. Derrick *et. al.*, *Measurement of the diffractive structure function in deep elastic scattering at HERA*, *Z. Phys.* **C68** (1995) 569–584, [[hep-ex/9505010](#)]. 1, 3, 16
- [38] **ZEUS** Collaboration, J. Breitweg *et. al.*, *Measurement of the diffractive structure function $F_2^{D_4}$ at HERA*, *Eur. Phys. J.* **C1** (1998) 81–96, [[hep-ex/9709021](#)]. 1, 3, 16, 4.3
- [39] **H1** Collaboration, C. Adloff *et. al.*, *Inclusive measurement of diffractive deep-inelastic ep scattering*, *Z. Phys.* **C76** (1997) 613–629, [[hep-ex/9708016](#)]. 1, 3, 4.3, 14

- [40] **ZEUS** Collaboration, J. Breitweg *et. al.*, *Measurement of the diffractive cross section in deep inelastic scattering using ZEUS 1994 data*, *Eur. Phys. J.* **C6** (1999) 43–66, [[hep-ex/9807010](#)]. 1, 3, 16, 4.4
- [41] **ZEUS** Collaboration, . S. Chekanov, *Deep inelastic inclusive and diffractive scattering at Q^2 values from 25 to 320 GeV^2 with the ZEUS forward plug calorimeter*, *Nucl. Phys.* **B800** (2008) 1–76, [[0802.3017](#)]. 1, 1, 3, 16, 4.4, 4.4, 15, 5, 16
- [42] E. Iancu, K. Itakura, and S. Munier, *Saturation and BFKL dynamics in the HERA data at small x* , *Phys. Lett.* **B590** (2004) 199–208, [[hep-ph/0310338](#)]. 1
- [43] A. H. Mueller and D. N. Triantafyllopoulos, *The energy dependence of the saturation momentum*, *Nucl. Phys.* **B640** (2002) 331–350, [[hep-ph/0205167](#)]. 1
- [44] D. N. Triantafyllopoulos, *The energy dependence of the saturation momentum from RG improved BFKL evolution*, *Nucl. Phys.* **B648** (2003) 293–316, [[hep-ph/0209121](#)]. 1
- [45] J. L. Albacete, N. Armesto, J. G. Milhano, and C. A. Salgado, *Non-linear QCD meets data: A global analysis of lepton- proton scattering with running coupling BK evolution*, *Phys. Rev.* **D80** (2009) 034031, [[0902.1112](#)]. 1, 2.3.3, 8, 3.2, 5, 5, 17, 5, 5, 6
- [46] J. L. Albacete, N. Armesto, J. G. Milhano, and C. A. Salgado, *A Global Analysis of DIS Data at Small- x with Running Coupling BK Evolution*, [0906.2721](#). 1, 3.2, 6
- [47] H. Weigert, *A compact introduction to evolution at small x and the color glass condensate*, *Nucl. Phys.* **A783** (2007) 165–172. 1
- [48] H. Weigert, *Running coupling in small x evolution: A summary*, *Acta Phys. Polon. Supp.* **1** (2008) 541–544. 1
- [49] H. Weigert, J. Kuokkanen, and K. Rummukainen, *Small x evolution in the CGC beyond the total cross section: Accommodating diffraction and other restrictions on the final state*, *AIP Conf. Proc.* **1105** (2009) 394–397. 1
- [50] C. Marquet and H. Weigert, *New observables to test the Color Glass Condensate beyond the large- N_c limit*, *Nucl. Phys.* **A843** (2010) 68–97, [[1003.0813](#)]. 1, 2.1, 2.3.2, 2.3.3
- [51] Y. V. Kovchegov, J. Kuokkanen, K. Rummukainen, and H. Weigert, *Subleading- N_c corrections in non-linear small- x evolution*, *Nucl. Phys.* **A823** (2009) 47–82, [[0812.3238](#)]. 2.1, 2.3.1, 2.3.1, 2.3.1, 2.3.1, C
- [52] N. N. Nikolaev and B. G. Zakharov, *Colour transparency and scaling properties of nuclear shadowing in deep inelastic scattering*, *Z. Phys.* **C49** (1991) 607–618. 2.1
- [53] C. Marquet, *A unified description of diffractive deep inelastic scattering with saturation*, *Phys. Rev.* **D76** (2007) 094017, [[0706.2682](#)]. 2.1, 2.2, 4.1, 4.1, 4.3, 31
- [54] E. Gardi, J. Kuokkanen, K. Rummukainen, and H. Weigert, *Running coupling and power corrections in nonlinear evolution at the high-energy limit*, *Nucl. Phys.* **A784** (2007) 282–340, [[hep-ph/0609087](#)]. 2.3, 2.3.2, 2.3.2, 2.3.2, E

- [55] Y. V. Kovchegov and H. Weigert, *Triumvirate of running couplings in small- x evolution*, *Nucl. Phys.* **A784** (2007) 188–226, [[hep-ph/0609090](#)]. [2.3](#), [2.3.2](#), [2.3.2](#), [2.3.2](#)
- [56] I. Balitsky, *Quark contribution to the small- x evolution of color dipole*, *Phys. Rev.* **D75** (2007) 014001, [[hep-ph/0609105](#)]. [2.3](#), [2.3.2](#), [2.3.2](#), [2.3.2](#)
- [57] I. Balitsky and G. A. Chirilli, *Next-to-leading order evolution of color dipoles*, *Phys. Rev.* **D77** (2008) 014019, [[0710.4330](#)]. [2.3](#), [2.3.2](#)
- [58] K. Rummukainen and H. Weigert, *Universal features of JIMWLK and BK evolution at small x* , *Nucl. Phys.* **A739** (2004) 183–226, [[hep-ph/0309306](#)]. [2.3](#), [2.3.1](#), [2.3.3](#)
- [59] T. Lappi, *Gluon spectrum in the glasma from JIMWLK evolution*, [1105.5511](#). [2.3.1](#)
- [60] A. Kovner and U. A. Wiedemann, *Eikonal evolution and gluon radiation*, *Phys. Rev.* **D64** (2001) 114002, [[hep-ph/0106240](#)]. [2.3.1](#)
- [61] E. Gotsman, E. Levin, U. Maor, and E. Naftali, *A modified Balitsky-Kovchegov equation*, *Nucl. Phys.* **A750** (2005) 391–405, [[hep-ph/0411242](#)]. [2.3.2](#)
- [62] Y. V. Kovchegov and H. Weigert, *Quark loop contribution to BFKL evolution: Running coupling and leading- N_f NLO intercept*, *Nucl. Phys.* **A789** (2007) 260–284, [[hep-ph/0612071](#)]. [2.3.2](#), [D.2](#), [D.2](#)
- [63] D. V. Shirkov and I. L. Solovtsov, *Analytic model for the QCD running coupling with universal $\bar{\alpha}_s(0)$ value*, *Phys. Rev. Lett.* **79** (1997) 1209–1212, [[hep-ph/9704333](#)]. [2.3.2](#), [2.3.3](#), [34](#)
- [64] I. L. Solovtsov and D. V. Shirkov, *The analytic approach in quantum chromodynamics*, *Theor. Math. Phys.* **120** (1999) 1220–1244, [[hep-ph/9909305](#)]. [2.3.2](#), [2.3.3](#), [34](#)
- [65] E. Iancu, K. Itakura, and L. McLerran, *Geometric scaling above the saturation scale*, *Nucl. Phys.* **A708** (2002) 327–352, [[hep-ph/0203137](#)]. [2.3.3](#)
- [66] **H1** Collaboration, C. Adloff *et. al.*, *Inclusive $D0$ and D^{*+} production in neutral current deep inelastic $e p$ scattering at HERA*, *Z. Phys.* **C72** (1996) 593–605, [[hep-ex/9607012](#)]. [11](#)
- [67] **H1** Collaboration, A. Aktas *et. al.*, *Measurement of $F2(c\bar{c})$ and $F2(b\bar{b})$ at high Q^2 using the H1 vertex detector at HERA*, *Eur. Phys. J.* **C40** (2005) 349–359, [[hep-ex/0411046](#)]. [11](#)
- [68] **H1** Collaboration, A. Aktas *et. al.*, *Measurement of $F2(c \text{ anti-}c)$ and $F2(b \text{ anti-}b)$ at low Q^{*2} and x using the H1 vertex detector at HERA*, *Eur. Phys. J.* **C45** (2006) 23–33, [[hep-ex/0507081](#)]. [11](#)
- [69] M. Wusthoff, *Large rapidity gap events in deep inelastic scattering*, *Phys. Rev.* **D56** (1997) 4311–4321, [[hep-ph/9702201](#)]. [4.1](#), [D.1](#), [31](#)
- [70] K. J. Golec-Biernat and M. Wusthoff, *Saturation effects in deep inelastic scattering at low Q^2 and its implications on diffraction*, *Phys. Rev.* **D59** (1999) 014017, [[hep-ph/9807513](#)]. [4.1](#)
- [71] K. Golec-Biernat and A. Łuszczak, *Dipole model analysis of the newest diffractive deep inelastic scattering data*, *Phys. Rev.* **D79** (2009) 114010, [[0812.3090](#)]. [4.1](#), [31](#)

- [72] S. Munier, A. Stasto, and A. H. Mueller, *Impact parameter dependent S matrix for dipole proton scattering from diffractive meson electroproduction*, *Nucl.Phys.* **B603** (2001) 427–445, [[hep-ph/0102291](#)]. 13
- [73] **ZEUS** Collaboration, S. Chekanov *et. al.*, *Study of deep inelastic inclusive and diffractive scattering with the ZEUS forward plug calorimeter*, *Nucl. Phys.* **B713** (2005) 3–80, [[hep-ex/0501060](#)]. 20
- [74] M. Braun, *Structure function of the nucleus in the perturbative QCD with $N_c \rightarrow \infty$ (BFKL pomeron fan diagrams)*, *Eur. Phys. J.* **C16** (2000) 337–347, [[hep-ph/0001268](#)]. 33
- [75] N. Armesto and M. A. Braun, *Parton densities and dipole cross-sections at small x in large nuclei*, *Eur. Phys. J.* **C20** (2001) 517–522, [[hep-ph/0104038](#)]. 33
- [76] K. Kutak and A. M. Stasto, *Unintegrated gluon distribution from modified BK equation*, *Eur. Phys. J.* **C41** (2005) 343–351, [[hep-ph/0408117](#)]. 33